

Superresolution of *GOES-16* ABI Bands to a Common High Resolution with a Convolutional Neural Network

CHARLES H. WHITE^①,^a IMME EBERT-UPHOFF,^{a,b} JOHN M. HAYNES,^a AND YOO-JEONG NOH^a

^a *Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado*

^b *Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, Colorado*

(Manuscript received 1 August 2023, in final form 21 November 2023, accepted 6 February 2024)

ABSTRACT: Superresolution is the general task of artificially increasing the spatial resolution of an image. The recent surge in machine learning (ML) research has yielded many promising ML-based approaches for performing single-image superresolution including applications to satellite remote sensing. We develop a convolutional neural network (CNN) to superresolve the 1- and 2-km bands on the GOES-R series Advanced Baseline Imager (ABI) to a common high resolution of 0.5 km. Access to 0.5-km imagery from ABI band 2 enables the CNN to realistically sharpen lower-resolution bands without significant blurring. We first train the CNN on a proxy task, which allows us to only use ABI imagery, namely, degrading the resolution of ABI bands and training the CNN to restore the original imagery. Comparisons at reduced resolution and at full resolution with *Landsat-8/Landsat-9* observations illustrate that the CNN produces images with realistic high-frequency detail that is not present in a bicubic interpolation baseline. Estimating all ABI bands at 0.5-km resolution allows for more easily combining information across bands without reconciling differences in spatial resolution. However, more analysis is needed to determine impacts on derived products or multispectral imagery that use superresolved bands. This approach is extensible to other remote sensing instruments that have bands with different spatial resolutions and requires only a small amount of data and knowledge of each channel's modulation transfer function.

SIGNIFICANCE STATEMENT: Satellite remote sensing instruments often have bands with different spatial resolutions. This work shows that we can artificially increase the resolution of some lower-resolution bands by taking advantage of the texture of higher-resolution bands on the *GOES-16* ABI instrument using a convolutional neural network. This may help reconcile differences in spatial resolution when combining information across bands, but future analysis is needed to precisely determine impacts on derived products that might use superresolved bands.

KEYWORDS: Remote sensing; Satellite observations; Artificial intelligence; Deep learning; Neural networks

1. Introduction

Superresolution (SR) is the general task of artificially increasing the spatial resolution of an image. While there have been many proposed techniques for superresolving images (Farsiu et al. 2004; Freeman et al. 2002; Glasner et al. 2009), an increasingly large number of recent approaches rely specifically on convolutional neural networks (CNNs; Yang et al. 2019). SR, in general, is an ill-posed problem since many possible high-resolution images exist for a given low-resolution image. This makes SR an inherently challenging task. As a result, CNNs and other machine learning approaches that minimize pixelwise error often yield blurry or overly smooth solutions (Wang et al. 2004). This, among other reasons, has led to the development of generative adversarial network (GAN)-based approaches for SR such as SRGAN (Ledig et al. 2017). SRGAN utilizes a second discriminator model that differentiates between superresolved and real images. In addition, it minimizes a content loss based on the intermediate activations of a separate pretrained CNN. This often yields an image with high-frequency detail more similar to ground truth but with higher pixelwise error (Ledig et al. 2017). Other

recent approaches have used diffusion models for superresolution (Li et al. 2022) that iteratively refine images by reversing a process that gradually adds noise, effectively enhancing image details at each step.

Deep learning SR techniques have recently seen wide use in the atmospheric and Earth sciences. Vandal et al. (2017) used a modified stacked SRCNN (Dong et al. 2015) to downscale precipitation data. Stengel et al. (2020) superresolved 50-times-higher resolution wind speed and 25-times-higher resolution solar irradiance fields from climate models. This is done using a two-stage process with two consecutive SRGAN-based models that first superresolve to an intermediate medium resolution and then again to the target high resolution. Geiss and Hardin (2020) explored the use of dense U-Net (Ronneberger et al. 2015; He et al. 2015) for superresolution of ground-based radar plan position indicator scans and found that they outperform commonly used techniques such as Lanczos and bicubic interpolation. Geiss and Hardin (2023) recently presented an approach to enforce invertibility in superresolution models, which is a property often desirable in many Earth science applications including remote sensing. In this context, invertibility ensures that the predicted high-resolution image can be used to exactly reproduce the original low-resolution image.

The vast majority of satellite remote sensing superresolution applications focus on low-Earth orbit (LEO) missions with spatial resolutions on the order of 1–100 m (Masi et al. 2016;

Corresponding author: Charles H. White, charles.white@colostate.edu

TABLE 1. The band number, central wavelength, and spatial resolution of the 16 channels present on ABI. Spatial resolutions listed here are representative for the subsatellite point, and they increase with larger viewing angles. Wavelength ranges and descriptions are quoted from [Schmit et al. \(2017\)](#) for *GOES-16* ABI specifically.

ABI band	Central wavelength (μm)	Wavelength range (μm)	Spatial resolution (km)	Description
Band 1	0.47	0.45–0.49	1	Blue
Band 2	0.64	0.60–0.68	0.5	Red
Band 3	0.86	0.847–0.882	1	Vegetation
Band 4	1.37	1.366–1.380	2	Cirrus
Band 5	1.6	1.59–1.63	1	Snow/ice
Band 6	2.2	2.22–2.27	2	Cloud particle size
Band 7	3.9	3.80–3.99	2	Shortwave window
Band 8	6.2	5.79–6.59	2	Upper-level water vapor
Band 9	6.9	6.72–7.14	2	Midlevel water vapor
Band 10	7.3	7.24–7.43	2	Lower/midlevel water vapor
Band 11	8.4	8.23–8.66	2	Cloud-top phase
Band 12	9.6	9.42–9.80	2	Ozone
Band 13	10.3	10.18–10.48	2	Clean longwave window
Band 14	11.2	10.82–11.60	2	Longwave window
Band 15	12.3	11.83–12.75	2	Dirty longwave window
Band 16	13.3	12.99–13.56	2	CO ₂

[Lanaras et al. 2018](#); [Scarpa et al. 2018](#); [Gargiulo et al. 2019](#); [Armannsson et al. 2021](#)) and revisit times on the order of multiple days. Fewer efforts have focused on geostationary (GEO) imagers with coarser spatial resolutions on the order of 1–10 km. One example is [Deneke et al. \(2021\)](#) which demonstrated that a relatively simple linear model can increase the spatial resolution of cloud property retrievals by using a high-resolution visible channel on a GEO imager. A clear advantage of GEO imagers over LEO imagers is the relatively high temporal resolution often making them more suitable for time-sensitive applications such as characterizing transient atmospheric or meteorological features. The Advanced Baseline Imager (ABI; [Schmit et al. 2017](#)) aboard the Geostationary Operational Environmental Satellite (GOES) R series satellites, for example, currently has a full-disk temporal resolution of 10 min and features a high resolution of 0.5 km for a single visible channel (band 2; 0.64 μm) and 2-km resolution for infrared wavelengths (see [Table 1](#) for a full list). Alternatively, the Sentinel-2 Multispectral Imager (MSI; [Drusch et al. 2012](#)), an instrument frequently used in SR applications ([Lanaras et al. 2018](#); [Wu et al. 2023](#); [Salgueiro Romero et al. 2020](#)), has a revisit time of 10 days and a spatial resolution of 10–60 m depending on the channel.

In this work, we superresolve the 15 lower-resolution channels out of the 16 channels present on ABI. Specifically, we artificially increase the resolution of the three 1-km channels and twelve 2-km channels on ABI to a common high resolution of 0.5 km. This constitutes a 2-times (2 \times) and 4-times (4 \times) increase in spatial resolution, respectively. Superresolving the lower-resolution channels could alleviate the need to reconcile differences in spatial resolution in the development of derived quantitative products (such as cloud, aerosol, or surface properties) or multispectral imagery from ABI. Our proposed approach leverages spatial texture in the higher-resolution channels to guide the estimation of subpixel radiance distribution in the lower-resolution channels. For example, low-resolution channels may contain cloud edges, shadows, and land and water boundaries. Band 2, in

some cases, provides information on the exact placement of these boundaries at high resolution. The presence of the 0.5-km channel on ABI allows the model to produce realistic texture in the super-resolved imagery during the day when visible reflectances are available. This is done without the use of adversarial training and without the added complexity and opacity of utilizing a separate pretrained model as is done in SRGAN. Additionally, it does not require multiple forward passes at inference time as is needed in current diffusion-based approaches.

Leveraging the texture of a high-resolution channel means that this work shares similarities with the objective of pan-sharpening ([Javan et al. 2021](#)). Pan-sharpening involves the use of a panchromatic band, which is a band with high spatial resolution but relatively wide spectral response. Spatial variability from the panchromatic band is used to sharpen other bands with low spatial resolution, but with narrower (and often overlapping) spectral response. In some respects, pan-sharpening could be considered a special case of SR where a panchromatic band is used to aid in superresolving the lower spatial resolution bands. In our case, the sole 0.5-km band on ABI is not necessarily panchromatic and its bandwidth does not overlap with other channels, but general features like cloud edges, shadows, and land and water boundaries are also observable in other ABI bands, which could help inform sub-pixel radiance distributions. Pan-sharpening, like SR, has seen many recent successful applications using CNN architectures (see [Tsagkatakis et al. 2019](#) for a review). Our approach uses a similar CNN architecture (with some modifications) as Deep Sentinel-2 (DSen2; [Lanaras et al. 2018](#)), which is a model trained to pan-sharpen Sentinel-2 imagery.

In this analysis, we train the CNN model on a proxy task using a synthetic lower-resolution dataset to learn a mapping that estimates the 1-km ABI channels at a 2 \times higher resolution and the 2-km channels at a 4 \times higher resolution. We perform both quantitative and qualitative evaluations in the reduced-resolution domain and compare the CNN to results from bicubic interpolation. Then, we use the CNN trained on

TABLE 2. The band number, approximate wavelength range, sensor, and spatial resolution of the channels present on *Landsat-8* and *Landsat-9*. Wavelength ranges quoted here are for *Landsat-8* (Barsi et al. 2014) and differ only slightly for *Landsat-9*.

OLI/TIRS bands	Wavelength range (μm)	Sensor	Spatial resolution (m)
Band 1	0.435–0.451	OLI	30
Band 2	0.452–0.512	OLI	30
Band 3	0.533–0.590	OLI	30
Band 4	0.636–0.673	OLI	30
Band 5	0.851–0.879	OLI	30
Band 6	1.566–1.651	OLI	30
Band 7	2.107–2.294	OLI	30
Band 8	0.503–0.676	OLI	15
Band 9	1.363–1.384	OLI	30
Band 10	10.60–11.19	TIRS	100
Band 11	11.50–12.51	TIRS	100

the lower-resolution task to the desired task of superresolving the original ABI bands to 0.5-km resolution. We show examples of predicted 0.5-km imagery from the CNN. In most typical SR and pan-sharpening applications, it is difficult to quantitatively compare estimated imagery at the target resolution due to a lack of ground-truth observations. We attempt to address this deficiency by collocating superresolved ABI imagery with National Aeronautics and Space Administration (NASA) and U.S. Geological Survey (USGS) *Landsat-8*/*Landsat-9* observations (Irons et al. 2012; Barsi et al. 2014; Wulder et al. 2019) that are made natively at much higher spatial resolution. Then, we quantify how well the CNN estimates fine-scale texture that is observed in Landsat imagery and compare it to bicubic interpolation. Last, we end with a discussion of the limitations of this approach, potential improvements for specific applications, and areas of future work.

2. Data

ABI is a GEO imager present on the four GOES-R series satellites. It measures radiance at 16 channels with a nadir spatial resolution of 0.5 km for band 2; 1 km for bands 1, 3, and 5; and 2 km for all other bands (Table 1). The temporal resolution of full-disk imagery varies with the scan mode used but is 10 min for the time period used in this work. All images from ABI in this analysis come from *GOES-16* in GOES-East position at a central longitude of 75.2°W. The ABI images used for training and evaluation of the neural network are selected from the first day of each month in 2021 at 1600, 1730, and 1900 UTC resulting in 36 full-disk images each consisting of 16 channels. The full disk is almost completely illuminated at 1730 UTC with the edges of the terminator visible on the western and eastern sides at 1600 and 1900 UTC. Our training set is created from images collected from 1 January 2021 through 1 October 2021. The validation set is created from the three images on 1 November 2021. The testing set is created from the three images collected on 1 December 2021.

Landsat-8 and *Landsat-9* are polar orbiting sun-synchronous satellites with 16-day repeat cycles. *Landsat-8* instruments include the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). OLI has 9 bands at visible and near-infrared wavelengths at a nadir spatial resolution of 30 m with

the exception of the panchromatic band (band 8) with a resolution of 15 m. TIRS has two bands near the infrared window region at 10.9 and 12.0 μm with a nadir spatial resolution of 100 m (Table 2). *Landsat-9* instruments (OLI-2 and TIRS-2) are improved versions of OLI and TIRS with improved radiometric resolution but are otherwise similar.

The Landsat data included in this work are not used during training and are only used to evaluate the superresolved images from the neural network at 0.5-km resolution after it has been trained. We collect 191 tiles from *Landsat-8* and *Landsat-9* centered around a 300-km radius from the *GOES-16* subsatellite point located at 0°, 75.2°W. These tiles are collected between 1 May 2022 and 1 August 2022. Additional *GOES-16* ABI data are obtained at the Landsat overpass times in order to collocate observations between the two instruments.

3. Methods

a. Constructing a supervised proxy task using the Wald protocol

Supervised machine learning problems require labeled data in order to train models. Since 0.5-km observations from ABI for all 16 channels only exist for band 2, we instead train our model on a proxy task that uses reduced-resolution imagery. This process consists of creating synthetic lower-resolution imagery and training a model to recreate the original native-resolution imagery. In the pan-sharpening literature, this methodology is commonly referred to as the Wald protocol (Wald 2002). Once trained, the model can be applied to native-resolution imagery to obtain the superresolved output. This protocol is illustrated in Fig. 1.

Our objective is to increase the resolution of the 2- and 1-km channels to 0.5-km resolution, which is a 4-times and 2-times increase in resolution, respectively. This means that in order to simulate a reduced-resolution training dataset, one must reduce the resolution of all 16 ABI channels to a 4-times-lower resolution. Thus, we degrade the 2-km ABI bands to an 8-km resolution, 1-km bands to 4 km, and the sole 0.5-km band to 2 km. The CNN is then trained to use the 8-, 4-, and 2-km inputs to produce imagery at a common resolution of 2 km. Once trained,

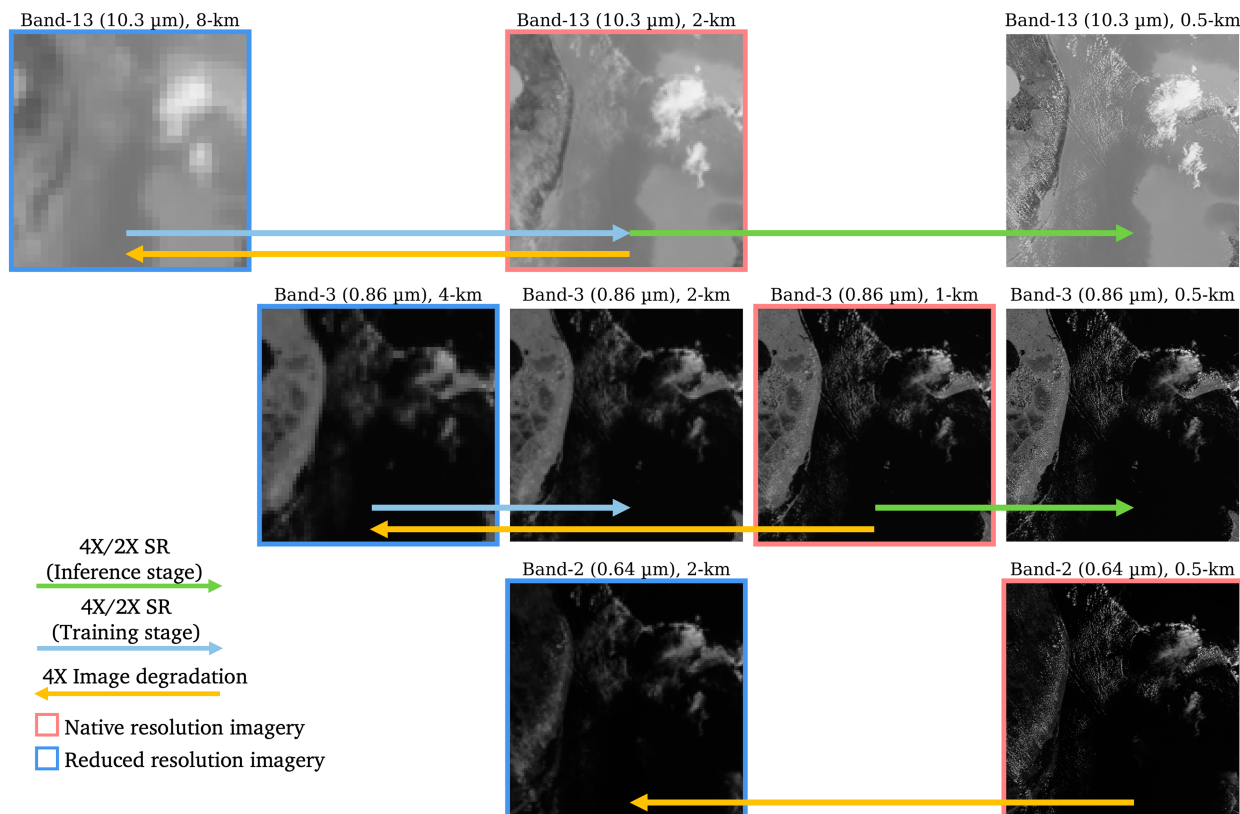


FIG. 1. A conceptual illustration of the different resolutions used during the training and inference stages for the (top) 12 native 2-km bands, (middle) three native 1-km bands, and (bottom) single native 0.5-km band from ABI. Images with salmon-colored borders represent the native resolution. Images bordered in blue are their counterparts at reduced resolution that are used as inputs when training the CNN model.

the model can use 2-, 1-, and 0.5-km inputs to superresolve all channels to a common resolution of 0.5 km.

Training at a reduced resolution, of course, involves a key assumption about these data. We effectively assume some amount of scale invariance between the transformations learned by the CNN for the synthetic reduced-resolution dataset and the observations at their native resolution. In other words, it is assumed that mapping pixels of an 8-km resolution image to 2-km resolution is the same transformation as mapping pixels of a 2-km resolution image to 0.5-km resolution. This assumption often goes untested in SR and pan-sharpening applications since imagery at the target resolution does not typically exist. We attempt to evaluate this assumption later in this analysis by comparing the texture of superresolved ABI imagery with *Landsat-8* and *Landsat-9*. [Armannsson et al. \(2021\)](#) evaluated this assumption in a similar way for Sentinel-2 pan-sharpening methods by using coarsened 5-m Airborne Visible Near-infrared Spectrometer (AVIRIS; [Green et al. 1998](#)) imagery. [Lanaras et al. \(2018\)](#) evaluated this assumption by applying their trained model across different scales and found a small increase in error when doing so.

The CNN we train here seeks to learn the inverse of the image degradation procedure used to create the synthetic reduced-resolution dataset. Therefore, it is important to consider

the subtleties of accurately simulating hypothetical ABI imagery at a $4\times$ lower spatial resolution. Specifically, one needs to account for the fact that satellite imagers do not have a perfect spatial response. For example, a point source contained within an imager pixel can contribute to the radiance observed in neighboring pixels (and vice versa). Quantifying the spatial response is critical for applications centered around bright fine-scale features such as characterizing properties of fires ([Calle et al. 2009](#)). The modulation transfer function (MTF; [Boreman 2001](#)) and the point spread function (PSF) are two ways of describing this characteristic of imaging systems. In our case, when creating the synthetic reduced-resolution dataset, one needs to account for the spatial response when downsampling to the reduced resolution.

Following the Wald protocol, we simulate the spatial response when creating the reduced-resolution dataset by blurring the image prior to downsampling. In practice, we achieve this by convolving the native-resolution imagery with an image filter defined by the PSF for each ABI channel. We assume that the PSF can be represented by a two-dimensional Gaussian distribution with a zero mean and a standard deviation calculated from known MTF values in Eq. (1), where σ represents the standard deviation of the PSF and s represents the scale ratio between images:

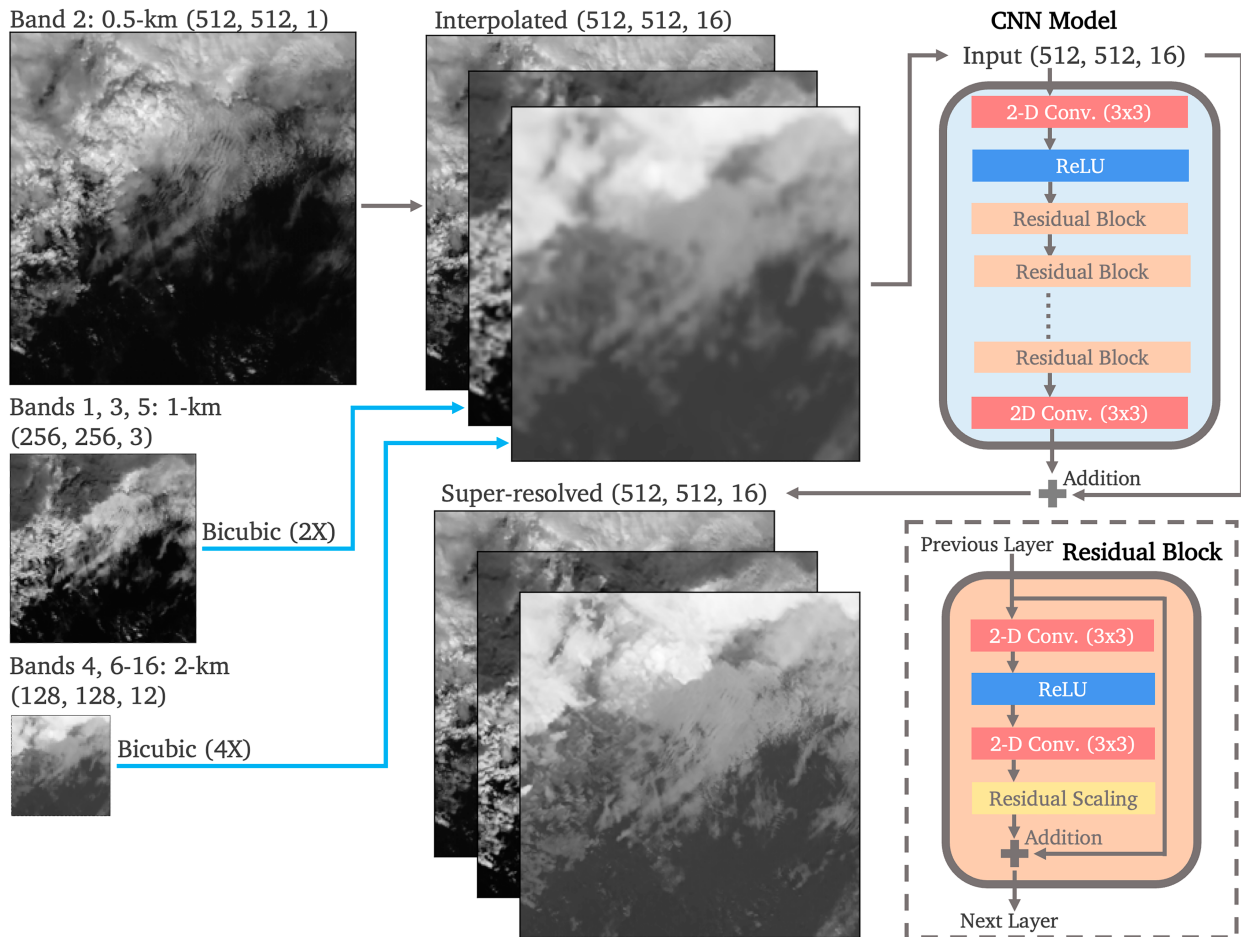


FIG. 2. Conceptual diagram of the CNN architecture used in this work. Lower-resolution channels are (top center) interpolated to 0.5 km using bicubic interpolation and (top right) passed through the CNN, which includes (bottom right) several residual blocks. (bottom center) The output of the CNN is added elementwise to the 15 interpolated bands to yield the superresolved imagery with nothing being added to band 2. See text for full description.

$$\sigma = s \left[\frac{-2\ln(\text{MTF})}{\pi^2} \right]^{1/2}. \quad (1)$$

In Eq. (1), we specifically use the value of the MTF at the Nyquist frequency. These values are obtained from the MTF estimates already made in Wu and Schmit (2019) for GOES-16 ABI. Without blurring the image with this kernel, the resulting low-resolution imagery would be too sharp. Thus, not accounting for an instrument's spatial response would likely yield a trained model that undersharpenes the native-resolution imagery.

b. Neural network architecture

The CNN architecture we use closely follows that of Lanaras et al. (2018). It consists of two-dimensional (2D) convolution layers, rectified linear unit (ReLU) activations, and residual scaling layers. A conceptual diagram of the CNN is shown in Fig. 2. Before the first convolutional layer, the resolution of each channel is increased to match the dimensions of band 2. This is done using a $2\times$ bicubic interpolation for bands 1, 3, and

5 and a $4\times$ bicubic interpolation for band 4 and bands 6–16. The 16-channel image is then passed to the first convolutional layer with a set number of 3×3 filters (N_F) followed by a ReLU activation. The output of the ReLU activation is passed through a number of residual blocks (N_{RB}), with N_F filters in each convolutional layer. After the last residual block, a final convolutional layer outputs 15 channels, which are then added elementwise to the original bicubically interpolated imagery. The skip connection between the inputs of the CNN and the output means that the model is trained to estimate a positive or negative adjustment to the pixels in the interpolated images given as input. The terms N_F and N_{RB} are both hyperparameters that can be tuned for an application-specific balance of performance and computational expense.

The residual blocks have a fairly simple design and follow those used in enhanced deep superresolution (EDSR) networks (Lim et al. 2017) and other pan-sharpening applications (Lanaras et al. 2018; Wu et al. 2020). They consist of a 2D convolution, ReLU activation, a second 2D convolution, and residual scaling (Szegedy et al. 2016) that multiplies the

activations by a scale factor of 0.1. The output of the residual scaling is then added to the input of the residual block. Lim et al. (2017) noted that removing batch normalization (Ioffe and Szegedy 2015) from residual blocks and adding residual scaling increased performance on superresolution tasks and improved stability during training. We similarly found that for this application, removing batch normalization and adding residual scaling allow the CNN to converge faster during training, reach a lower loss, and be more numerically stable.

The receptive field in this model is relatively small compared to other image-to-image translation applications that use U-Net-based architectures (Ronneberger et al. 2015; Zhou et al. 2018; Huang et al. 2020). As noted in Lanaras et al. (2018), much of the information relevant to sharpening the lower-resolution channels is likely local in space to any individual pixel. Thus, a larger receptive field may not be necessary. During the development of this approach, we noted that using various types of U-Nets with much larger receptive fields but with a similar total number of trainable parameters typically yielded imagery with roughly 1.25–1.5-times-larger mean absolute error (MAE). We did not perform a thorough hyperparameter search with the U-Net models, so a more formal comparison is not included here.

c. Training details and data preparation

Full-disk images for *GOES-16* ABI band 2 at 0.5-km resolution have dimensions of 20968×20968 pixels and are much too large to be processed in GPU memory all at once. For training purposes, we divide the full-disk images into 512×512 tiles at 0.5-km resolution. Tiles with more than 30% of pixels outside the full disk are discarded. No other restrictions based on viewing or solar geometry are applied to the tiles. Pixels outside the full disk are assigned a sample weight of zero and are not counted in the loss function or evaluation metrics. This yields 1424 tiles for each full-disk image. Following the training, validation, and testing split indicated previously, this results in 42 720 tiles for training, 4272 for validation, and 4272 for testing. The image degradation procedure described previously is used to create reduced-resolution tiles at 8-, 4-, and 2-km resolutions with resulting dimensions of 32×32 , 64×64 , and 128×128 pixels, respectively. The model is trained to sharpen this imagery to a uniform 2-km resolution (128×128). During the development of this model, we noted that models trained on only 30% of the images used here performed very similarly in terms of MAE on our validation set. Since this model is trained using synthetic data, our training set could feasibly be increased to a much larger size. However, we expect further increases in performance from doing so to be relatively small.

The CNN operates on scaled radiance (Schmit and Gunshor 2021) read directly from the ABI files without converting to true radiance. Scaled radiances are integers with values ranging from a minimum of 0 to a valid maximum of 16 382 or less depending on the channel. We standardize the scaled radiance by subtracting the mean and dividing it by the standard deviation for each channel independently. This makes converting to

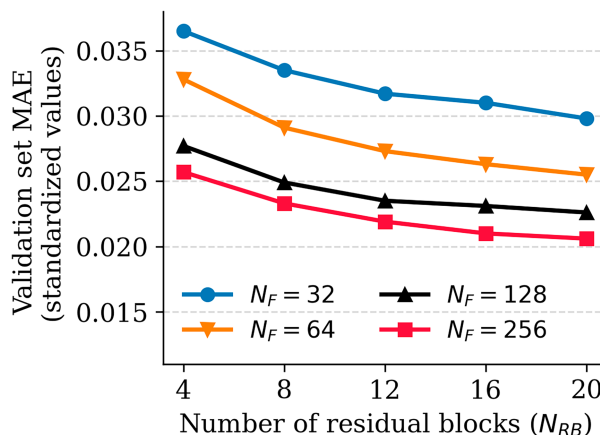


FIG. 3. Results of a hyperparameter search over the number of residual blocks (N_{RB}) and the number of filters per convolutional layer (N_F) for the reduced-resolution proxy task.

true radiance an unnecessary step since the resulting standardized values would be the same due to the linear relationship between scaled radiance and radiance. The conversion to radiance, reflectance, or brightness temperature can instead be performed on the CNN's output after multiplying by the standard deviation and adding the mean for each channel. See Schmit and Gunshor (2021) for details about these conversions.

MAE is used as the loss function to train the CNN. Other loss functions were considered including adding an additional term to penalize differences in high-pass-filtered imagery similar to Lu and Chen (2019). The addition of this component of the loss did not appear to significantly change the quality of the estimated imagery for this specific application so it is not included here. The CNN is trained with a batch size of 24 images with an initial learning rate of 5×10^{-4} . The loss is calculated on the entire validation set after every 1000 batches. The learning rate is reduced by a factor of 10 to a minimum of 5×10^{-6} when no improvement in the validation loss is observed for 2000 batches. Training is stopped when no improvement is observed for 4000 batches. Models typically finished training at around 35 000 batches.

Figure 3 shows how the performance of the neural network changes according to N_{RB} and N_F for this reduced-resolution proxy task. The validation set MAE was calculated at the reduced resolution and averaged across all bands except band 2. MAE decreased monotonically with increasing N_{RB} and increasing N_F over the range of values that we tested. Figure 3 implies that a model with $N_{RB} = 20$ and $N_F = 256$ has the lowest error. However, in the following analysis, we will use the model with $N_{RB} = 12$ and $N_F = 256$. Although there are models with better performance, this model was selected as a compromise between performance and computational expense with near-real-time applications in mind. A single model with $N_{RB} = 12$ and $N_F = 256$ trains in roughly 3 h on a single NVIDIA RTX A6000 GPU. Our implementation of this model can produce a superresolved full-disk image in roughly 2.5 min and a superresolved mesoscale sector in roughly 3 s.

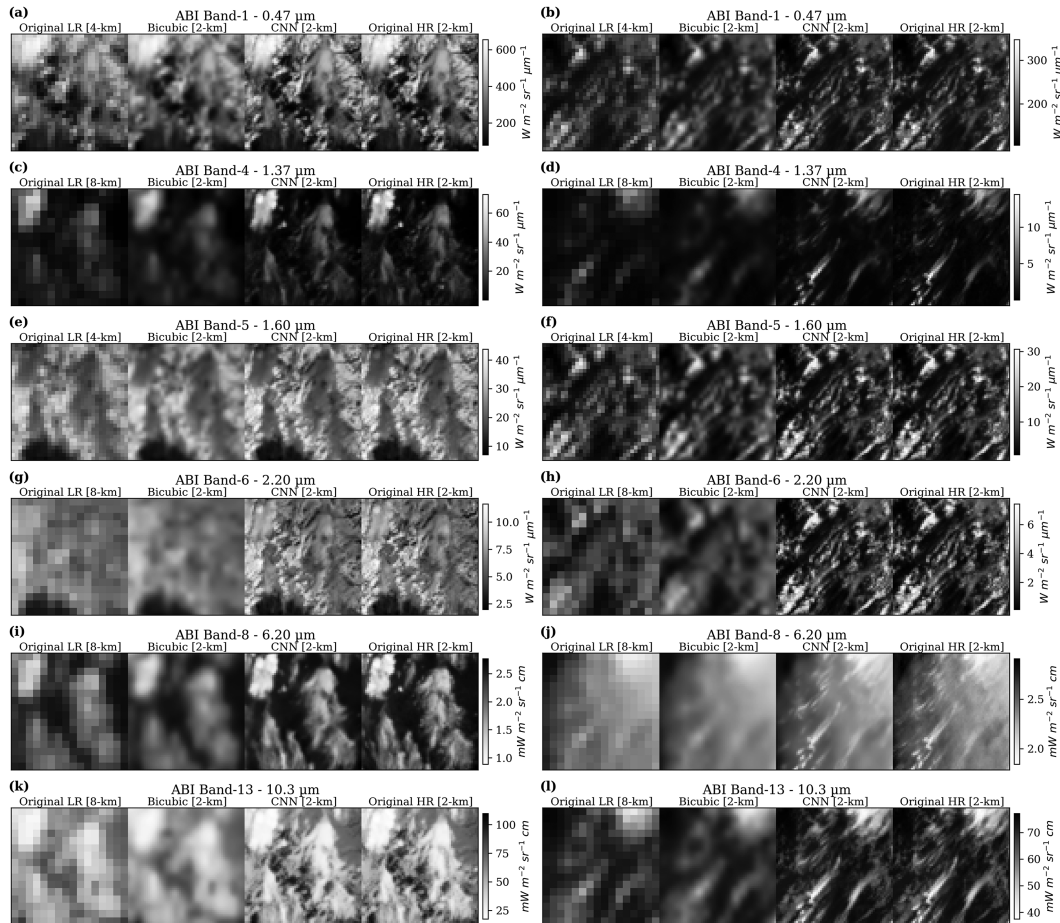


FIG. 4. Examples of original LR inputs to the CNN, bicubic interpolation, CNN output, and target HR imagery. Shown are selected images from bands (a),(b) 1, (c),(d) 4, (e),(f) 5, (g),(h) 6, (i),(j) 8, and (k),(l) 13 for two sample cases in the left and right columns.

4. Results

a. Comparisons at reduced resolution

We begin our analysis of the superresolved imagery with qualitative and quantitative comparisons to reference imagery in the reduced-resolution domain (superresolving to 2-km resolution from 8 and 4 km). Where appropriate, we compare to bicubic interpolation—a simple baseline approach that does not take into account subpixel variability observed in band 2 and does not utilize spectral variability in any way. Since the CNN uses interpolated imagery as input, the difference in performance between the CNN and bicubic interpolation may be interpreted as the added value of using the CNN to super-resolve ABI imagery.

1) VISUAL IMAGERY COMPARISON

Figure 4 shows a qualitative comparison of the reduced-resolution imagery at 8 and 4 km (LR), bicubic interpolation to a uniform resolution of 2 km, the CNN output at 2 km, and the original high-resolution imagery at 2 km (HR) for a few selected bands for two sample cases. In all cases, as expected,

bicubic interpolation tends to produce very fuzzy images. As a result, many fine-scale features present in the original high-resolution imagery appear blurry or nonexistent in the bicubic interpolated imagery. Band 4 in Fig. 4c and band 6 in Fig. 4h are particularly striking examples of this. As expected, ABI channels with only a $2\times$ increase in resolution are less blurred by bicubic interpolation.

The CNN output at 2 km overall appears significantly sharper and is more similar to the target high-resolution imagery. In many cases, features present in the CNN output are not recognizable in the original LR images but are easily identified in the corresponding HR images. This implies that the high-resolution observations in band 2 heavily influence the subpixel variability present in the CNN output. Figures 4a and 4b show examples of fine-scale clouds in the CNN output that are not present in the bicubic interpolation but match well with features in the original HR imagery. In contrast to bicubic interpolation, the CNN more frequently estimates radiances larger than those in the original LR imagery especially in band 4 and improves the appearance of finer-scale cloud features. In the superresolved imagery, it is much easier

TABLE 3. MAE and RMSE of bicubic interpolation and the superresolved imagery from the CNN calculated on the reduced-resolution dataset. Ratios near 0 indicate large improvement in the CNN predictions relative to bicubic interpolation. Ratios near 1 indicate little to no improvement. MAE and RMSE are expressed in units of $\text{W m}^{-2} \text{sr}^{-1} \mu\text{m}^{-1}$ for bands 1–6 and in units of $\text{mW m}^{-2} \text{sr}^{-1} \text{cm}$ for bands 7–16.

ABI band	Central wavelength (μm)	MAE bicubic	MAE CNN	MAE ratio (CNN/bicubic)	RMSE bicubic	RMSE CNN	RMSE ratio (CNN/bicubic)
Band 1	0.47	6.996	0.448	0.064	13.416	0.751	0.056
Band 3	0.86	4.43	0.497	0.112	7.963	1.086	0.136
Band 4	1.37	0.422	0.145	0.344	1.226	0.375	0.306
Band 5	1.6	0.779	0.13	0.167	1.31	0.243	0.185
Band 6	2.2	0.374	0.064	0.171	0.594	0.106	0.178
Band 7	3.9	0.03	0.01	0.333	0.05	0.019	0.380
Band 8	6.2	0.029	0.014	0.483	0.062	0.025	0.403
Band 9	6.9	0.087	0.038	0.437	0.191	0.072	0.377
Band 10	7.3	0.181	0.071	0.392	0.402	0.142	0.353
Band 11	8.4	1.278	0.35	0.274	2.229	0.642	0.288
Band 12	9.6	0.861	0.225	0.261	1.586	0.424	0.267
Band 13	10.3	1.887	0.504	0.267	3.308	0.934	0.282
Band 14	11.2	2.058	0.557	0.271	3.645	1.048	0.288
Band 15	12.3	1.991	0.574	0.288	3.622	1.094	0.302
Band 16	13.3	1.314	0.413	0.314	2.559	0.789	0.308

to delineate between upper- and lower-level clouds in band 6 in Fig. 4g. Despite appearing sharper than bicubic imagery, the CNN still appears to underestimate extreme values. Texture in band 8 (Figs. 4i,j) and finer-scale features in band 4 (Figs. 4c,d) are also smoother compared to the original HR images. This illustrates the limitations of using only a $0.6\text{-}\mu\text{m}$ band to inform the spatial texture in other bands sensitive to different spectral properties of the surface, clouds, and atmosphere.

2) RESULTS FOR PIXELWISE ERROR: MAE AND RMSE

Next, we evaluate the performance of the CNN by calculating the MAE and the root-mean-square error (RMSE; Table 3). These metrics are calculated after converting the output of the CNN and bicubic interpolation to radiance. Note that radiance is integrated over wavelength for ABI bands 1–6 with units of $\text{W m}^{-2} \text{sr}^{-1} \mu\text{m}^{-1}$ and integrated over wavenumber for ABI bands 7–16 with units of $\text{mW m}^{-2} \text{sr}^{-1} \text{cm}$. Since the range of typical radiance values varies widely across different channels, we also calculate the ratio of CNN and bicubic metrics. Ratios near 1.0 imply little benefit through using the CNN, and values close to zero imply very large improvement. For all channels, we find very substantial improvement with the CNN compared to bicubic interpolation for both MAE and RMSE. Specifically, for MAE, we observe that the CNN has 6.4% of the error that bicubic interpolation has for band 1 ($0.47 \mu\text{m}$). At worst, we see that the CNN has 48.3% of the MAE that bicubic interpolation has for band 8 ($6.20 \mu\text{m}$).

There are a few key points made by the ratios in Table 3. First, we see that bands 1, 3, and 5 have much lower ratios than the other bands. One major reason for this is that only $2\times$ SR is required for these bands. Another reason is that these bands, along with band 6, typically covary strongly with band 2 and often have similar texture, which also makes it an easier task to superresolve them. All other bands require $4\times$ SR, and many

are sensitive to different properties of the atmosphere compared to band 2. For example, bands 8–10 (6.2 , 6.9 , and $7.3 \mu\text{m}$) are relatively sensitive to water vapor absorption. This makes these bands difficult to superresolve since they share little similarity to band 2. However, band 2 could still be informative in determining boundaries of upper-level clouds as seen in band 8 in Fig. 4i.

3) RESULTS FOR SPATIAL STRUCTURE: SSIM

To evaluate the estimated spatial texture from the bicubic interpolation and the CNN output, we use the structural similarity index measure (SSIM) that was developed specifically to assess the degradation of structural information in images in an effort to mimic human visual perception (Wang et al. 2004). SSIM is calculated on a sliding window throughout an image and has a maximum ideal value of 1 indicating very similar images, a minimum value of -1 indicating anticorrelation, and a central value of 0 indicating no similarity. Example SSIM values for bicubic interpolation and the CNN output can be seen in Fig. 5 for band 13 and band 6. The relatively large difference between the bicubic interpolation SSIM and the CNN's SSIM represents the differences in spatial texture in these example images. This difference is particularly evident in Fig. 5d where many bright, fine-scale features are smoothed over in the interpolated images.

The mean SSIM is calculated across the entire testing set separately for each ABI channel and is shown in Fig. 6. The relatively large SSIM values for bicubic interpolation in the water vapor absorption bands (6.20 , 6.90 , and $7.3 \mu\text{m}$) could represent the relatively smooth texture typically observed at these wavelengths. The stark difference between bicubic interpolation and the CNN at band 6 ($2.20 \mu\text{m}$) is likely a combination of multiple factors. This channel requires a $4\times$ increase in spatial resolution but typically covaries strongly with band 2. The CNN is able to take this relationship into account, but

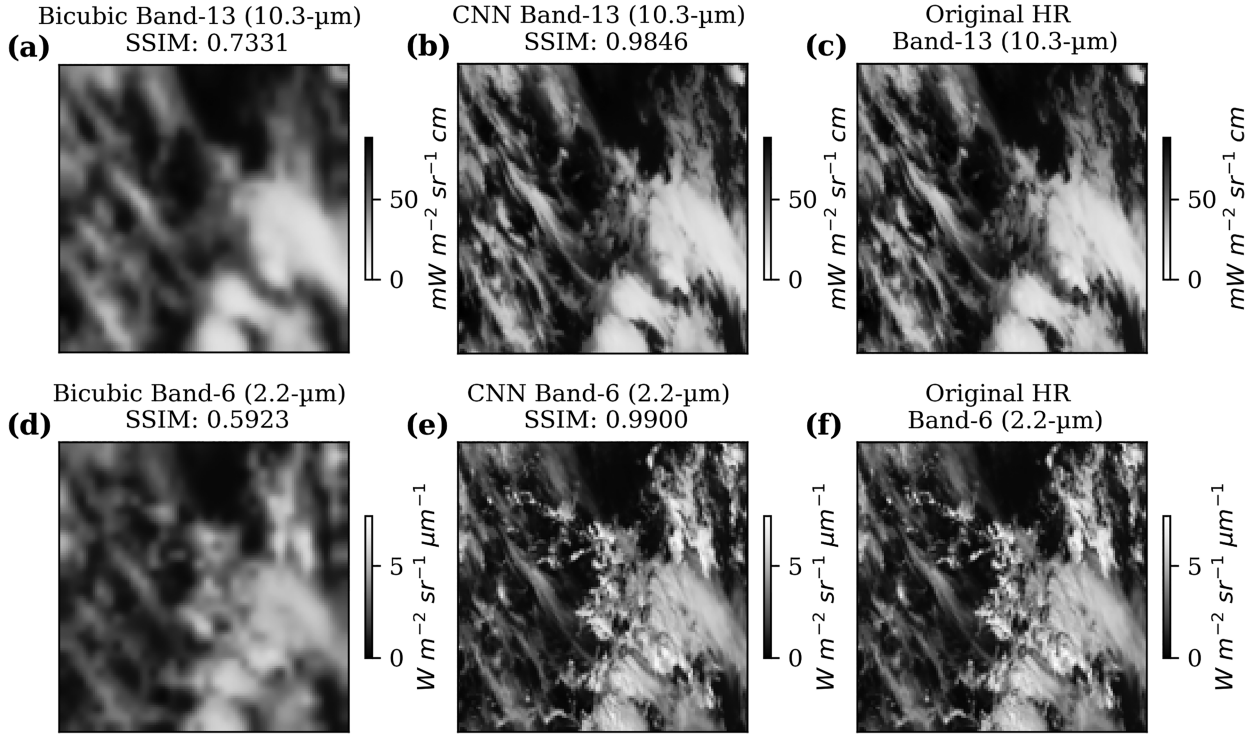


FIG. 5. Example SSIM values calculated for (a),(d) bicubic interpolation and (b),(e) superresolved images relative to the original HR images, in the reduced-resolution dataset. (c),(f) Also shown are the original HR band-13 and band-6 images for these examples.

bicubic interpolation does not. For all channels, the CNN outperforms bicubic interpolation in terms of SSIM. Larger differences occur in the infrared channels at $3.90 \mu\text{m}$ and with wavelengths longer than $8.4 \mu\text{m}$, and the largest difference occurs with the $2.2\text{-}\mu\text{m}$ band. It is difficult to draw strong conclusions from SSIM alone when comparing improvement across channels. However, it is an additional piece of evidence to show the value added by the CNN compared to bicubic interpolation on a per-channel basis.

4) RESULTS FOR SPECTRAL SIMILARITY: SAM

Quantitative satellite products from ABI such as sea and land surface temperature, cloud detection, cloud and aerosol properties, volcanic ash detection, and many others rely more on spectral properties of the observations rather than spatial patterns. These spectral properties can consist of simple differences between brightness temperatures, ratios between reflectances, or more complicated spectral indices. Thus, it is useful to explore how well the superresolved imagery preserves relationships between channels and how this compares to the bicubic baseline approach. Evaluations of image processing methods often use spectral angle mapper (SAM; Kruse et al. 1993) as a metric to assess spectral distortion in estimated imagery. Originally designed as a classification algorithm, SAM values have since been used as a metric to describe the similarity between two spectra (Deborah et al. 2015). This is done by measuring the angle formed by two vectors representing the predicted and reference spectra. SAM

values describe the same property of images as cosine similarity (CS), another often-used metric in the remote sensing and computer vision literature. SAM and CS can be described and related by Eqs. (2) and (3), where \hat{y}_i is the scalar predicted value for the i th band, and y_i is the reference value for this band. Ideal values of SAM are smaller with a minimum of 0, and ideal values of CS are larger with a maximum of 1:

$$\text{CS} = \frac{\sum_{i=1}^{16} \hat{y}_i y_i}{\sqrt{\sum_{i=1}^{16} \hat{y}_i^2} \sqrt{\sum_{i=1}^{16} y_i^2}}, \quad \text{and} \quad (2)$$

$$\text{SAM} = \cos^{-1}(\text{CS}). \quad (3)$$

In the following, we report results using SAM. For each pixel in a pair of superresolved and original HR images, we calculate the SAM values between the two 16-dimensional vectors that represent the 16 bands in an ABI image. Larger SAM values, expressed in degrees, indicate larger differences in spectra and imply that relationships between channels may not be preserved in the predicted high-resolution imagery. ABI bands with a larger range of observed radiances can have a disproportionate influence on the direction the vectors point, so we calculate SAM on radiances normalized by their maximum observed value. However, we found similar results when calculated on radiances without normalization. The distributions of SAM values are shown in Fig. 7 for all images in

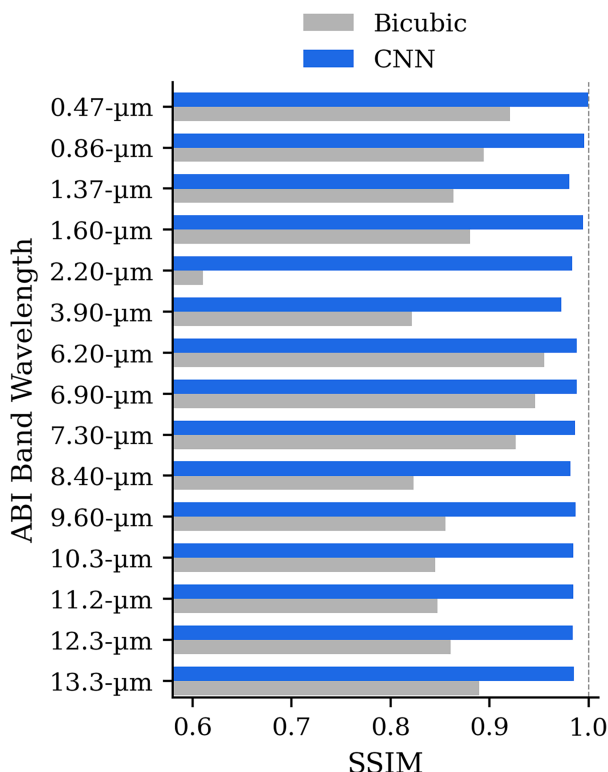


FIG. 6. SSIM values calculated on the testing dataset and averaged for each channel for bicubic interpolation (gray) and the CNN (blue) on the reduced-resolution dataset. SSIM for bicubic interpolation at $2.20\ \mu\text{m}$ is 0.62. The gray dashed line indicates the maximum ideal SSIM value of 1.0.

the testing dataset for both the CNN and bicubic interpolation. The CNN images have a much larger proportion of pixels with lower SAM values compared to the bicubic images. The mean SAM value is also almost 5 times larger for bicubic interpolation compared to the CNN. This indicates that the CNN is more likely to better preserve relationships across channels compared to the bicubic interpolation baseline.

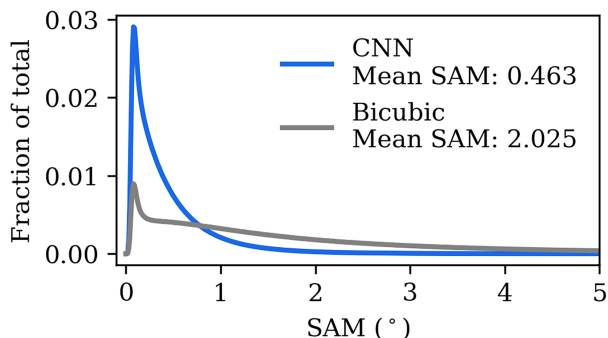


FIG. 7. Histograms of the SAM values for bicubic interpolation and the CNN estimates. The y axis represents the fraction of total pixels within each of the equally spaced bins.

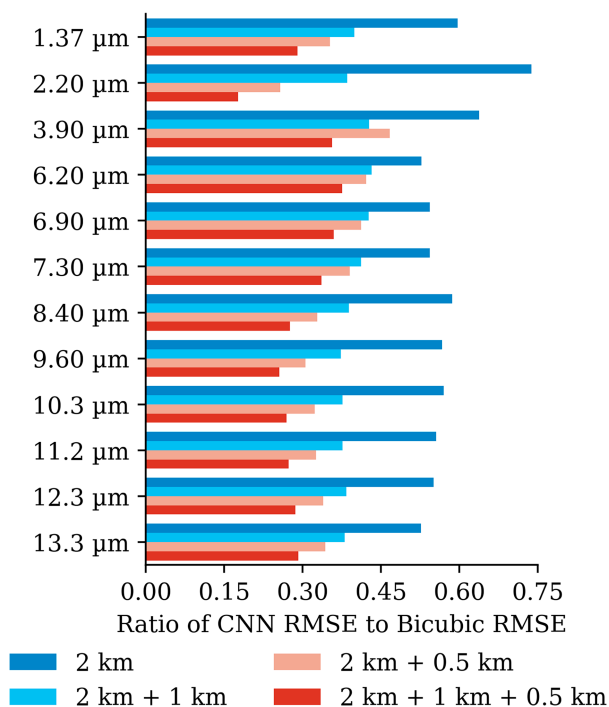


FIG. 8. Assessment of the impact of adding different sets of bands to the CNN on the superresolution of the 2-km bands.

5) ASSESSMENT OF THE VALUE OF THE HIGH-RESOLUTION BANDS

The model shown in the previous analyses appears to add substantial value to bicubic interpolation. Our hypothesis is that the higher-resolution bands aid the superresolution of the 2-km bands. In an effort to quantify the impact of the high-resolution information, we retrain four different models to only superresolve the 2-km channels. Each of these four models has access to a different combination of inputs according to their spatial resolution shown in Fig. 8. Reduction in the ratio of CNN RMSE to bicubic RMSE with the additional 1- or 0.5-km bands indicates improvement in CNN performance. These four models are all trained in the same way with $N_{\text{RB}} = 12$ and $N_F = 256$. Note that some of these ratios may differ from Table 3 since the prediction task here does not include the 1-km channels.

Figure 8 illustrates that without the 1- and 0.5-km bands, the CNN can superresolve the 2-km bands with 53%–74% of the RMSE of bicubic interpolation. With the addition of either the 1-km bands or the single 0.5-km band, the results improve substantially. For all but the $3.90\text{-}\mu\text{m}$ band, the single 0.5-km band offers more improvement than the three 1-km bands. As expected, in all bands, the lowest RMSE occurs when using all available high-resolution information to superresolve the 2-km channels. However, the amount of improvement that comes with the addition of the high-resolution information varies between bands. It is worth noting that our model is designed specifically to exploit the information available in the highest-resolution band by upsampling the lower-resolution channels and convolving the images at a native

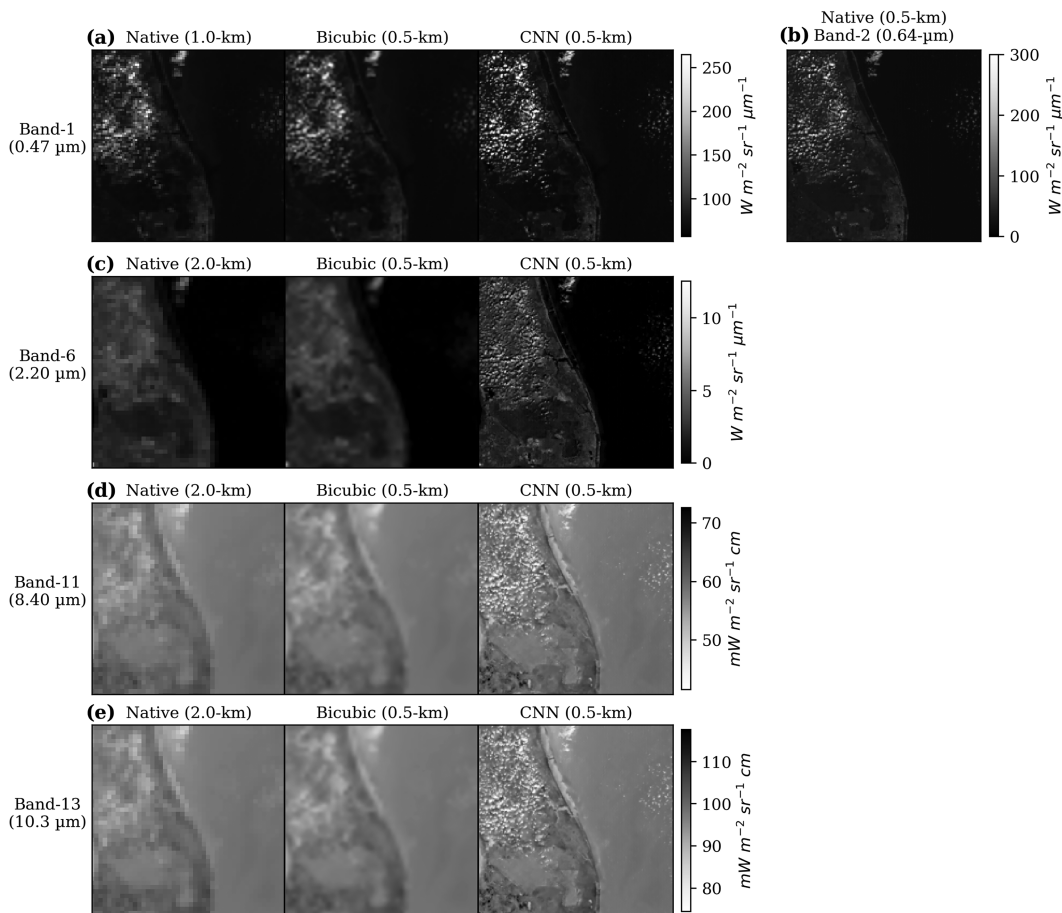


FIG. 9. (a),(c),(d),(e) Qualitative comparison of the native-resolution imagery, bicubic interpolation to 0.5 km, and CNN SR to 0.5 km from left to right in each subpanel. (b) The native 0.5-km band-2 imagery. Each subpanel shows a different band indicated on the left side. Note the differences in the color bar and units for each band. This scene has a central latitude and longitude of 27.14°N, 80.17°W and was acquired at 1600 UTC 1 Dec 2021.

resolution of 0.5 km. If only 2-km channels were available in practice, then there are other model architectures that may be a better choice.

b. Comparisons at full resolution

In the previous section, we evaluated the superresolved imagery from the CNN in the reduced-resolution domain, where we sharpen 8- and 4-km imagery to 2 km. We showed that the CNN adds value to bicubic interpolation for all channels through qualitative and quantitative comparisons measured by pixelwise error (MAE and RMSE), spatial texture (SSIM), and spectral similarity (SAM). Next, we turn to the full-resolution domain where the CNN superresolves 2- and 1-km imagery to a uniform 0.5-km resolution. These images are more difficult to evaluate due to the lack of ground-truth reference data at the target 0.5-km resolution.

1) FULL-RESOLUTION IMAGERY COMPARISON

We first start with a qualitative analysis of the predicted 0.5-km imagery for selected bands. Figure 9 shows bands 1, 6,

11, and 13. For each channel, we show native resolution, bicubic interpolation to 0.5 km, and the CNN output at 0.5-km resolution. Band 2 is also shown to illustrate the high-resolution information used in the CNN. For band 1 and band 6, we can observe much greater detail in the surface features and more well-defined low-level cloud cover over land. For band 6 specifically, we see low-level broken cloud cover over the ocean in the eastern portion of the image that is not easily distinguished in native-resolution band 6. This cloud cover seen in the CNN imagery appears realistic and is faintly visible in other bands such as in native-resolution band 1 and band 2. Bicubic interpolation does not show these features for band 6. Infrared wavelengths are also shown in band 11 and band 13 (8.4 and 10.3 μm). These infrared bands do not have much similarity with the 1- and 0.5-km channels making it difficult to evaluate the quality of this imagery from a qualitative perspective. However, there are more well-defined features present in the CNN output compared to the native-resolution image and the bicubic interpolation. The CNN shows similar cloud features both over land and over the ocean seen in band 2.

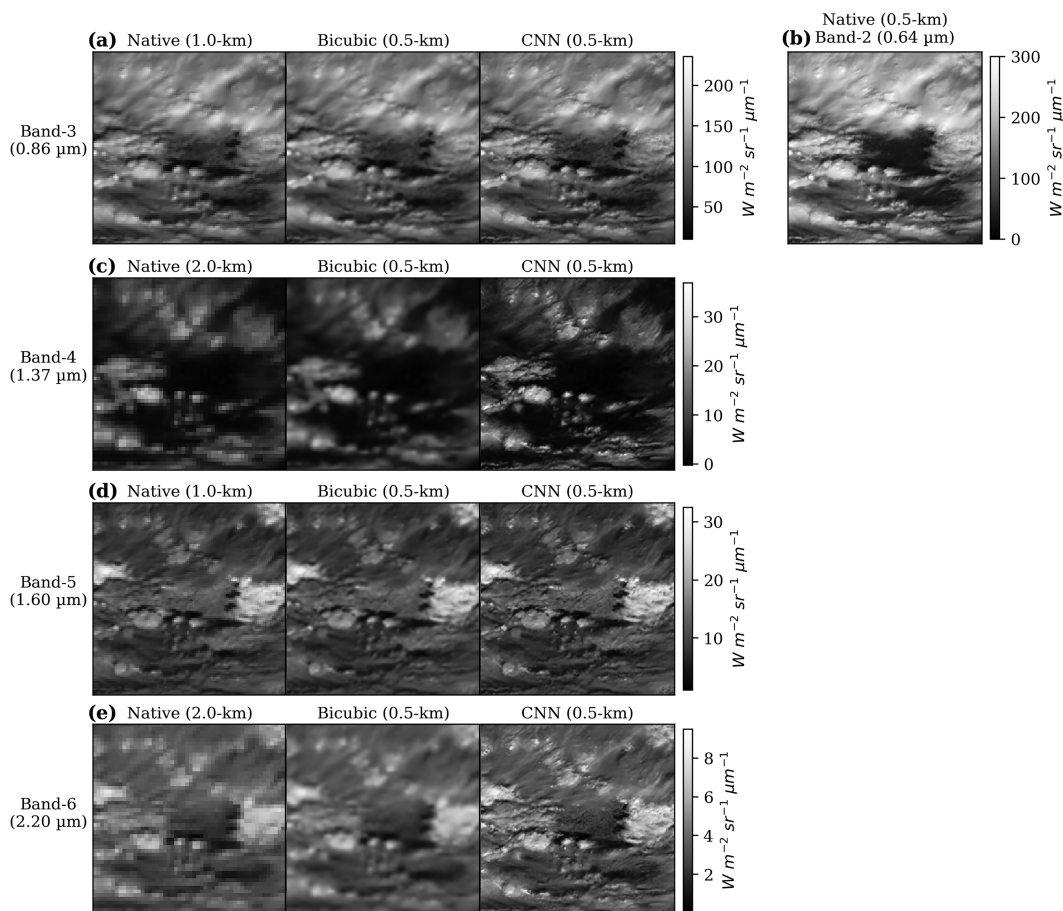


FIG. 10. As in Fig. 9, but with a central latitude and longitude of 36.98°N , 80.71°W . These images were acquired at 1600 UTC 1 Dec 2021.

The CNN also shows sharper gradients relative to the native-resolution imagery corresponding to surface features, and coastal boundaries in the infrared channels.

Next, we compare a selection of visible and near-infrared wavelengths sharpened to 0.5 km in Fig. 10. This scene shows more vertically developed clouds casting shadows on the land surface and contains both upper- and lower-level cloud cover. In all bands shown, the CNN estimates offer more detailed imagery compared to the native-resolution and bicubic imagery. Band 4 is somewhat unique among the visible/near-infrared channels shown here since it is mostly sensitive to upper-level clouds due to strong water vapor absorption. Band 5 and band 6 have some sensitivity to cloud phase and are typically darker (smaller radiance) for ice clouds and brighter (larger radiance) for liquid water clouds. Thus, it is possible that band 5 with a native resolution of 1 km, in addition to band 2, can play a role in sharpening the texture of band 4 for upper-level cirrus. This imagery, to a limited extent, illustrates the ability of the CNN to stay true to the general appearance of bands 5 and 6 despite the delineation between ice and water clouds being less apparent in band 2. Despite only being faintly visible in band 2, the fine-scale surface features in the center of these images appear well-resolved in the CNN imagery compared to bicubic interpolation.

Overall, the difference between bicubic interpolation and the CNN is less striking for bands 3 and 5 due to only a $2\times$ increase in spatial resolution.

To examine bands 7–10, which include short- and midwave infrared wavelengths, we choose a scene with a moderate amount of water vapor and a mix of high- and low-altitude clouds. Bands 8, 9, and 10 are particularly sensitive to water vapor absorption at the upper, middle, and lower levels of the atmosphere, respectively. These bands typically do not share much similarity in texture with band 2 except for dry atmospheres and high-altitude clouds. In Fig. 11, it is clear that the band-2 texture in the high-altitude cloud band in the middle of the images could be used to sharpen similar features in the water vapor bands. The low-level cloud feature in the upper right of the band-2 imagery is apparent in band 10. However, this feature becomes less pronounced as water vapor absorption increases in bands 8 and 9. A similar effect is observed with the cloud features in the bottom right of these images. The larger area of strong water vapor absorption in the upper left of these images is relatively homogeneous. This is a case where bicubic interpolation appears relatively accurate and where superresolving the image with a CNN adds little value. Strongly absorbing and spatially homogeneous water vapor features are likely one reason why the CNN has

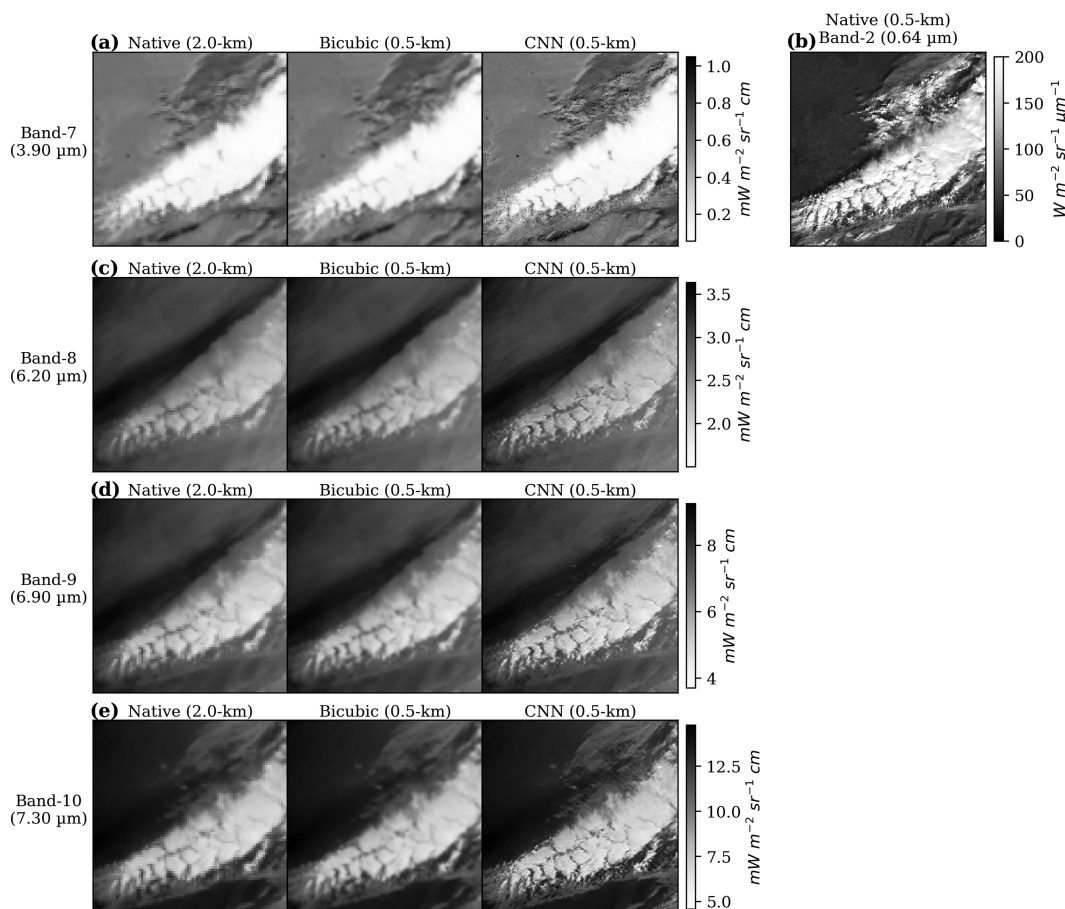


FIG. 11. As in Fig. 9, but with a central latitude and longitude of 36.41°N, 91.28°W. These images were acquired at 1600 UTC 1 Dec 2021.

48.3% of the MAE for band 8 compared to smaller percentages for bands 11–16 (Table 3 in the reduced-resolution dataset).

2) COMPARISONS TO LANDSAT IMAGERY

While the imagery in Figs. 9–11 gives some confidence in the quality of the superresolved imagery from the CNN, this alone is not enough to determine whether the superresolved details are realistic or correct. Next, we perform a quantitative evaluation of the superresolved 0.5-km imagery by comparing it to observations from *Landsat-8* and *Landsat-9* imagers that are made natively at 30- and 100-m resolution depending on the channel (Table 2). These data are collocated with full-disk ABI images collected at roughly the same time at ABI's subsatellite point. The median difference in acquisition time between instruments is 2 min but can be as large as 5 min. Only superresolved imagery from channels with roughly similar central wavelengths will be compared with Landsat in this section.

(i) An example of superresolved ABI and coarsened Landsat imagery

Figure 12 shows a portion of one of the 191 image matchups between the Landsat satellites and ABI obtained between

1 May 2022 and 1 August 2022. There are nine channels (including band 2) on ABI that have roughly similar central wavelengths present on the Landsat imagers (OLI/TIRS). Qualitatively, it appears that the CNN superresolved imagery has captured many fine-scale details that are also present in analogous OLI/TIRS channels. The CNN seems to appropriately sharpen the low-level broken cumulus that are present in the OLI/TIRS particularly in ABI band 1 and ABI band 6 (Figs. 12d,g). These features are blurred and, in some cases, appear as spatially homogeneous low-level clouds in the bicubic interpolated images.

The band-13 CNN imagery appears to be an improvement over bicubic interpolation overall, but there are some minor issues. The cirrus cloud edges visible in TIRS band 10 (Fig. 12k) appear slightly smoother than those seen in the ABI band-13 superresolved image (Fig. 12j). There are also a few areas where the texture of lower-level broken clouds visible in band 2 may have been added to the cirrus clouds in band-13 CNN output or is overemphasized in translucent clouds. When comparing to TIRS band 10, these features look much smoother, implying the CNN may have oversharpened ABI band 13 in some areas. This, again, may show a clear disadvantage in only using a visible

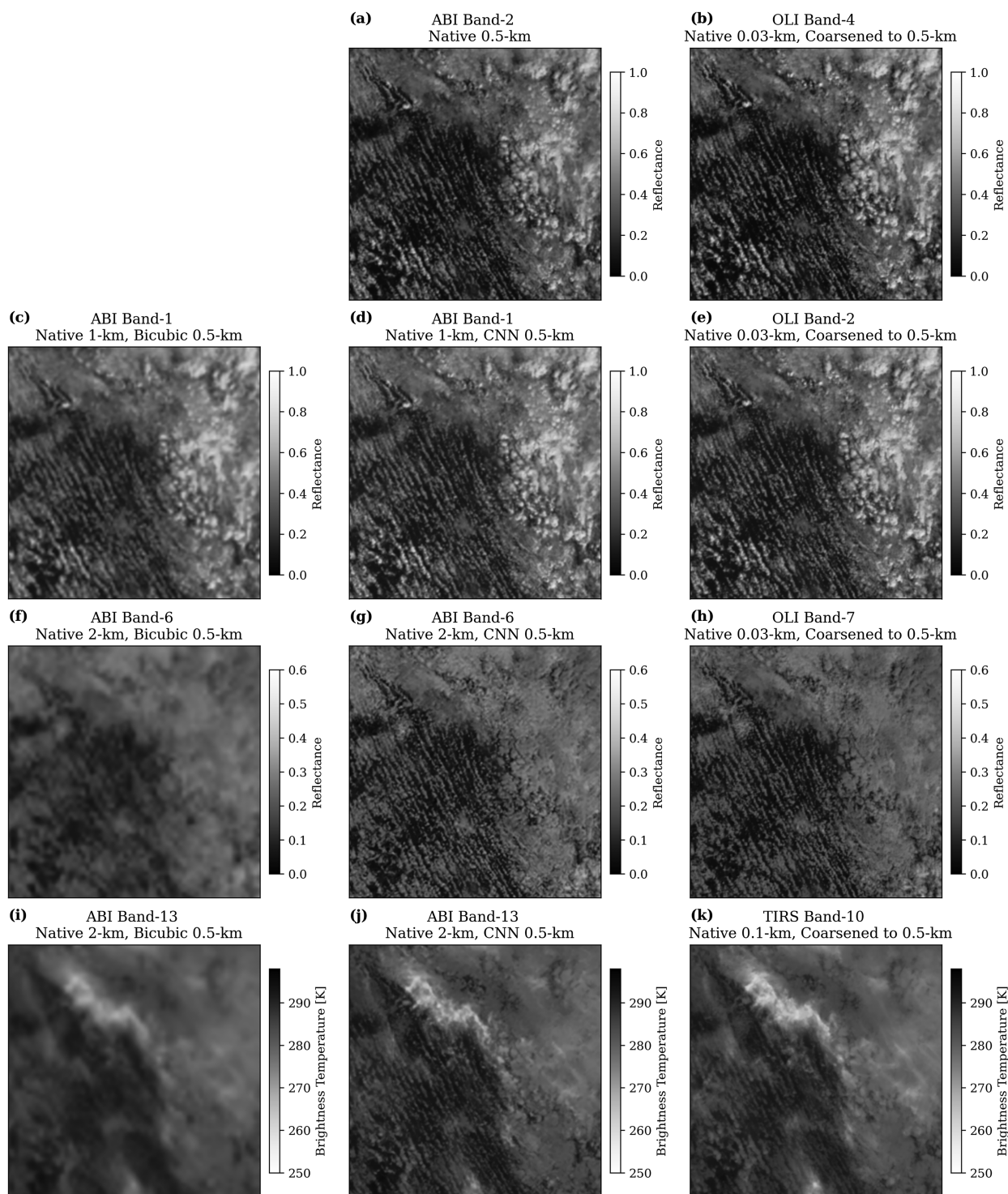


FIG. 12. An example of a collocated scene between *GOES-16* ABI and *Landsat-9* OLI/TIRS. (left) The bicubic interpolation to 0.5 km for bands (c) 1, (f) 6, and (i) 13. (center) Shown are (a) native-resolution band-2 and (d),(g),(j) superresolved imagery from the CNN. (right) Shown are (b),(e),(h),(k) OLI/TIRS imagery coarsened to a resolution of 0.5 km.

band to sharpen imagery at wavelengths sensitive to other radiative properties of the surface, clouds, and atmosphere. A relatively small part of this specific textural difference between ABI and TIRS could also be due to differences in spectral response at these wavelengths. Low-level cloud texture (absent any upper-level cirrus) and cloud boundaries appear to be better represented in the superresolved CNN imagery in all bands shown in Figs. 12d, 12g, and 12j.

Despite these images occurring within a median time difference of 2 min and with similar viewing geometry, significant care needs to be taken when comparing these two instruments quantitatively. At a resolution of 0.5 km and finer, even small time differences can allow for the advection of cloud features into neighboring pixels. This potential for slight spatial mismatching between ABI and Landsat favors smoother imagery when comparing individual collocated pixels along the edges of fine-scale cloud features. Differences in spectral response between the OLI/TIRS and ABI channels are an additional source of error that will weaken the overall correspondence between the measurements of these two sensors but will not necessarily favor one method over the other.

(ii) *Comparison of selected collocations with Landsat*

In an effort to reduce issues associated with the time difference between ABI and OLI/TIRS images, we focus the first portion of this analysis on collocated pixels where local spatial patterns match relatively well between both instruments. For this purpose, we use the native 0.5-km observations from ABI band 2 and 0.5-km coarsened imagery from OLI band 4. For every pixel in the images, we then calculate the Pearson correlation coefficient in a 3 pixel \times 3 pixel sliding window between these two channels. When the correlation coefficient is near 1.0, it implies good correspondence in the local spatial patterns of both images. Large positive correlations also imply, but do not guarantee, minimal issues with the movement of clouds between images. We select pixels in these images for comparison when their local correlations are larger than 0.90. Note that this selection criterion likely favors stationary surface features, images with smaller time differences, and lower-level cloud cover due to the increased likelihood of greater wind speeds at high altitudes.

After selecting pixels with similar local spatial patterns in the observations, we evaluate the quality of the superresolved ABI channels by comparing them to analogous channels on OLI/TIRS. Figure 13 shows the correlation between ABI and OLI/TIRS after applying two different high-pass filters. These values are shown for bicubic interpolation and for the CNN, each with a 3×3 filter [$\mathbf{G}_{3 \times 3}$; Eq. (4)] and a 5×5 filter [$\mathbf{G}_{5 \times 5}$; Eq. (5)]. We choose to compare high-pass-filtered imagery since, ideally, the value of the CNN processing comes from the additional high-frequency detail. These two high-pass filters are used to assess the high-frequency detail produced in the estimated imagery at slightly different scales. In channels with 2-km resolution, $\mathbf{G}_{3 \times 3}$ summarizes detail contained entirely within a native-resolution pixel. At the other extreme, $\mathbf{G}_{5 \times 5}$ summarizes

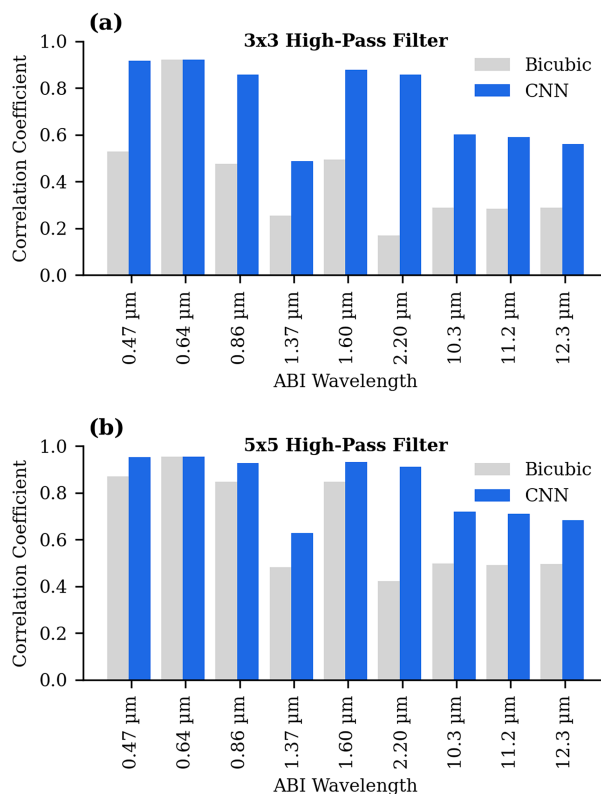


FIG. 13. Selected-collocation comparison: Pearson correlation coefficients between the estimated 0.5-km imagery for ABI and the corresponding coarsened OLI/TIRS observations for selected pixels after applying a (a) 3×3 and (b) 5×5 high-pass filter.

high-frequency information across multiple native-resolution pixels in the native 1-km channels:

$$\mathbf{G}_{3 \times 3} = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}, \quad \text{and} \quad (4)$$

$$\mathbf{G}_{5 \times 5} = \begin{bmatrix} -1 & -1 & -1 & -1 & -1 \\ -1 & 1 & 2 & 1 & -1 \\ -1 & 2 & 4 & 2 & -1 \\ -1 & 1 & 2 & 1 & -1 \\ -1 & -1 & -1 & -1 & -1 \end{bmatrix}. \quad (5)$$

The results from Fig. 13 illustrate that the CNN more reliably estimates fine-scale texture present in the coarsened 0.5-km Landsat imagery. The added value of using the CNN relative to bicubic interpolation varies with wavelength and the size of the filter. The channels with the highest correlations in local texture are at visible and near-infrared wavelengths except for the 1.37- μm channel. This is likely because these channels share more information with band 2, and in the case of the 0.47- and 0.86- μm channels, only require a $2 \times$ increase in resolution. The infrared channels have some of the lowest correlation with OLI/TIRS but still offer substantial improvement over bicubic interpolation. We expect that the relatively low correlations between ABI band 4 (1.37 μm) and the

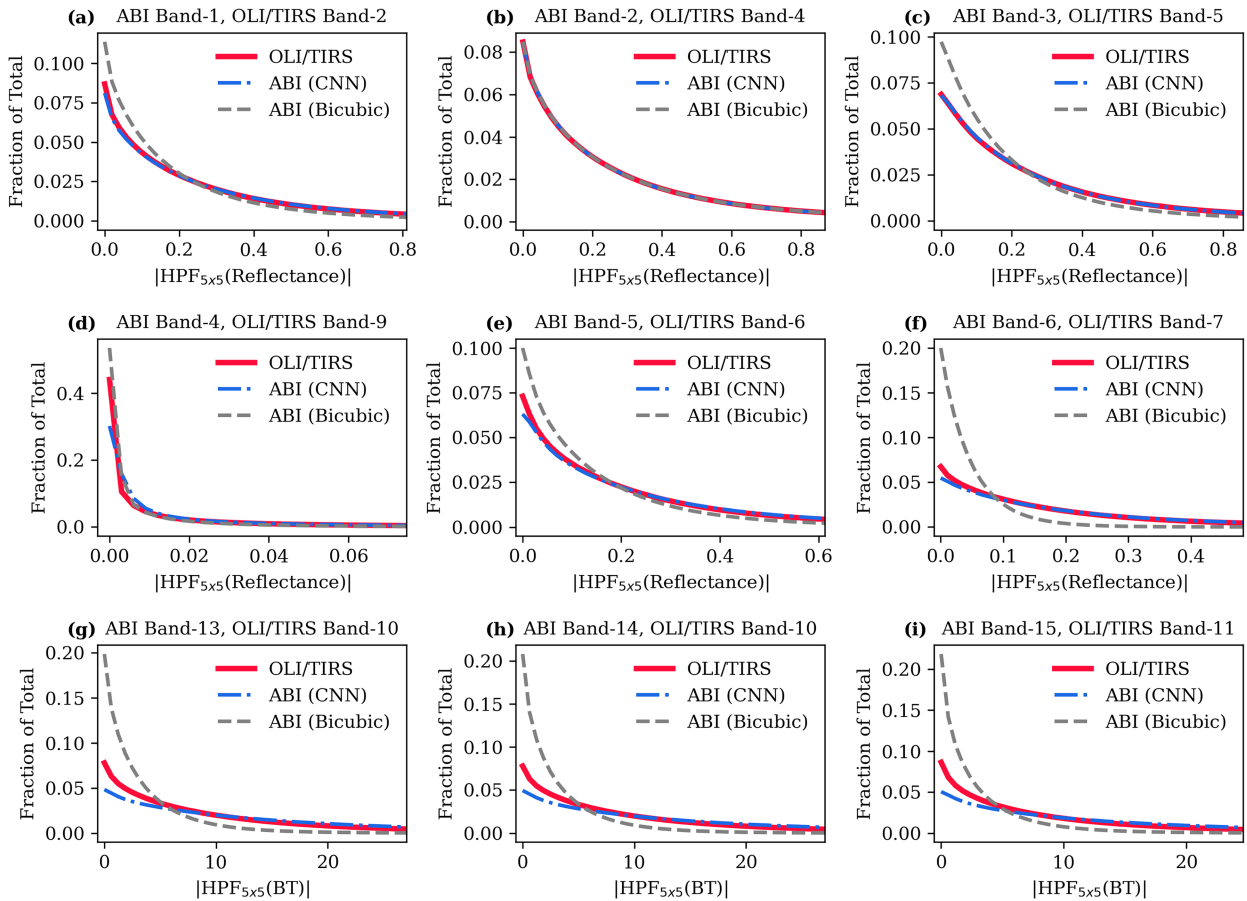


FIG. 14. (a)–(i) All-collocation histograms: Histograms of high-pass-filtered imagery using the 5×5 filter in Eq. (5) and taking all pixels into account. Each panel shows a different comparison of roughly comparable channels from ABI and OLI/TIRS. Note that the two ABI histograms in (b) are the same since band 2 has a native resolution of 0.5 km. The x axis limits for each subplot are defined by the 0th and 95th quantile of the Landsat OLI/TIRS distributions.

corresponding OLI/TIRS channel may be, in part, due to the time difference between images. The $1.37\text{-}\mu\text{m}$ channel is mostly sensitive to upper-level cloud features where high wind speeds are more likely to advect clouds larger distances over a given period of time.

(iii) Comparison across all collocations with Landsat

The requirement used above unfortunately limits the comparison to a sample size of 1 364 236, which is only 5.36% of the roughly 25.3 million total collocated observations. As an alternative approach for evaluating the quality of the CNN output, we compare the overall distributions of the values after convolving them with $\mathbf{G}_{5 \times 5}$ (Fig. 14). This approach reduces issues caused by the movement of clouds except when advected outside the edge of these images and allows us to consider all collocated observations. However, it ignores the exact spatial matching between pixels used in Fig. 13.

Figure 14 shows the distributions of high-pass-filtered channels from ABI after bicubic interpolation and CNN SR to 0.5 km, as well as similarly filtered OLI/TIRS imagery. Comparisons between the native 0.5-km ABI band 2 and OLI/TIRS

band 4 show very similar distributions. The small differences between them might be attributed to each instrument's spectral response among other minor factors. Comparing distributions for other channels reveals that the CNN superresolved imagery has a texture more similar to that of OLI/TIRS compared to bicubic interpolation. As expected, the CNN provides much more high-frequency detail shown by the closely aligned histograms of the CNN and OLI/TIRS. The bicubic interpolation is much smoother shown by the higher frequency of smaller gradients. We also note a consistent tendency for the CNN to over-sharpen the three infrared bands included in Fig. 13. These channels have a smaller proportion of pixels with low gradients and a larger proportion with larger gradients from the CNN. This is consistent with the idea that the CNN may be inserting texture from band 2 or other 1-km channels, which does not exist in the 2-km infrared channels.

We can quantify the differences in the distributions in Fig. 14 by calculating the one-dimensional Earth mover's distance (EMD). EMD can be succinctly described as the amount of work needed to transform one distribution into another. In two or more dimensions, it requires solving an optimal-transport

TABLE 4. One-dimensional EMD calculated for the histograms shown in Fig. 14. EMD values are divided by 100 to ease comparison. Ratios less than 1.0 in the rightmost column indicate channels where the CNN's distribution of high-pass-filtered values more closely aligns with OLI/TIRS.

ABI band	OLI/TIRS band	Bicubic EMD/100	CNN EMD/100	Ratio (CNN/bicubic)
Band 1	Band 2	2.466	0.346	0.140
Band 2	Band 4	0.040	0.040	1.000
Band 3	Band 5	2.596	0.067	0.026
Band 4	Band 9	13.82	5.253	0.380
Band 5	Band 6	3.917	0.957	0.244
Band 6	Band 7	11.99	1.937	0.162
Band 13	Band 10	0.180	0.094	0.521
Band 14	Band 10	0.181	0.095	0.526
Band 15	Band 11	0.181	0.113	0.624

problem. In one dimension, it reduces to Eq. (6), where F_1 and F_2 are cumulative distribution functions:

$$\text{EMD} = \int_0^q |F_1(x) - F_2(x)| dx. \quad (6)$$

Here, q is the largest value of the x axis in Fig. 14 that we integrate over which we set to the 99th quantile of the high-pass-filtered Landsat imagery. We calculated the EMD between the OLI/TIRS and the two ABI histograms after converting them into probability density functions with the results shown in Table 4. These EMD values and their ratios illustrate quantitatively that for every channel comparison with OLI/TIRS, the CNN has a more similar distribution of high-pass-filtered reflectances and brightness temperatures.

5. Discussion

Overall, the CNN offers substantial value over bicubic interpolation for nearly every channel and nearly every metric in this analysis, but the exact amount of improvement is channel dependent. The best results are seen for visible and near-infrared channels that are most similar to the native 0.5-km band 2. Specifically, the most reliable results are for bands 1, 3, 5, and 6, which all offer large improvements over bicubic interpolation in the reduced- and full-resolution domains and match the OLI/TIRS gradient distributions very well. Other bands, such as band 13, also show a large improvement in both the reduced- and full-resolution domains but have a slightly weaker correspondence to OLI/TIRS observations and show some issues with atmospheric features that band 2 is not sensitive to. ABI bands with strong water vapor absorption (bands 8–10) show improvement in the reduced-resolution domain through using the CNN and appear realistic upon qualitative inspection in the full-resolution domain, but do not have comparable channels on OLI/TIRS. We expect similar issues with strongly absorbing water vapor features that co-occur in space with unrelated features with high texture in ABI band 2.

One previously mentioned concern for quantitative products created using superresolved imagery is that spectral relationships between channels may not be preserved in the CNN's output. Many satellite products rely on differences or ratios between two or more channels. Therefore, it is essential

that relevant spectral relationships are maintained in the superresolved output. We are able to quantify the spectral distortion at reduced resolution, and our results illustrate that the CNN appears to have fewer issues with unrepresentative spectra compared to bicubic interpolation. However, we are not able to evaluate this in the full-resolution domain due to the unavailability of certain channels on OLI/TIRS and differences in spectral response functions. If there is a specific spectral relationship that is particularly important to a given application, one might consider a few strategies during model development for mitigating potential issues. One approach could include adding a term to the loss function that penalizes incorrectly estimating specific derived spectral features that might include relevant band differences or ratios. If there is a specific satellite product of interest, another alternative might be to superresolve the specific satellite product itself. In that case, the product could be added as an additional channel to the input and output. More general solutions might involve adding cosine similarity to the loss function to directly incentivize more realistic spectra (He et al. 2020).

The model we train here uses scaled radiance as input, but many derived satellite products operate on reflectances or brightness temperatures. This has no impact for the visible and near-infrared channels since there is a linear relationship between radiance and reflectance. However, there is a nonlinear relationship between radiance and brightness temperature. A small increase in radiance will change brightness temperatures more at low radiances compared to large radiances. Figure 15 shows this relationship for ABI band 13. This means that a 1-K error in brightness temperature is penalized by the model more heavily in warm scenes compared to cold scenes. If high accuracy is needed in cold scenes, such as the estimation of cloud properties for opaque high-altitude clouds, then it may be useful to train the model on brightness temperatures or scale the loss function according to an equivalent change in brightness temperature.

There are several other potential improvements that could be made to the model analyzed here. The model we used in this analysis has 12 residual blocks and 256 filters per convolutional layer. Figure 3 illustrates that further increases in either of these hyperparameters could yield some benefit, but likely with diminishing returns and increased computational expense.

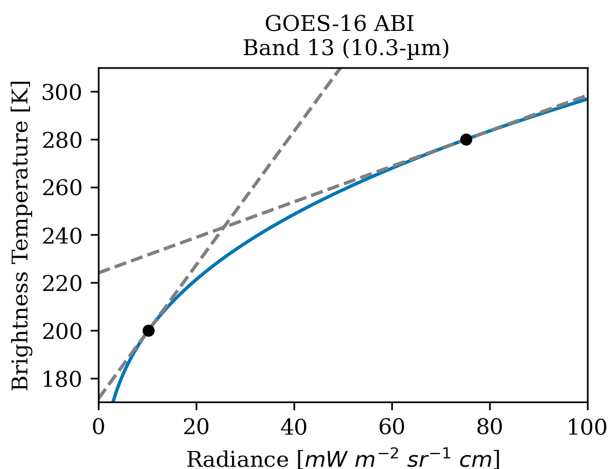


FIG. 15. Relationship of brightness temperature and radiance for *GOES-16* ABI band 13 (solid blue line). Gray dashed lines indicate the slope of the curve at two sample points indicated by black dots.

One advantage GEO imagers have over LEO imagers is the high temporal resolution, which we have not taken advantage of here. Several studies have shown the benefit of including temporal information in deep learning SR (Liu et al. 2022), suggesting that this may be one way of further improving ABI SR. Other architectural improvements could include applying high-pass-filtering layers to the input imagery (Gargiulo et al. 2019), adding attention layers (Liu et al. 2021), or using deeper densely connected residual blocks (Wang et al. 2019). As mentioned previously, Geiss and Hardin (2023) presented an approach that enforces invertibility between the superresolved and input images. While their approach, as presented, does not account for the spatial response of remote sensing instruments, it could be modified and included here to help ensure physical consistency between the original low-resolution and predicted high-resolution images. In this work, consistency between the superresolved output and the input images is only incentivized through the MAE loss function, which does not guarantee invertibility.

A clear limitation of this approach is that band-2 radiances are not available during the night. This means the model must learn the $4\times$ SR mapping for infrared channels without any higher-resolution information. Our testing dataset includes a small subset of images along the day–night terminator. Predictions in these images (not shown) for the infrared channels were, as expected, blurry and appeared to be only a very slight improvement over bicubic interpolation. Therefore, we do not expect the CNN to add much value beyond bicubic interpolation during the night.

This overall methodology is readily extensible to other remote sensing instruments given that they have channels with differing spatial resolutions. The Visible Infrared Imaging Radiometer Suite (VIIRS; Hillger et al. 2013) aboard the Suomi National Polar-orbiting Partnership (SNPP) and Joint Polar Satellite System (JPSS) satellites, for example, has 16 moderate-resolution channels at 750-m resolution and 5 imaging channels at 375-m resolution for visible, near-infrared, and infrared wavelengths. We might expect higher quality SR

of infrared channels for VIIRS compared to ABI due to the wider spectral coverage of the 375-m channels. Similarly, the presence of 375-m infrared channels on VIIRS likely allows for accurate nighttime SR unlike the presented approach for ABI.

There are also clear applications to future planned satellite missions such as the National Oceanic and Atmospheric Administration (NOAA) Geostationary Extended Observations (GeoXO) satellite system. While the final specifications of the GeoXO Imager (GXI) are not yet known, the current GXI requirements detail many upgrades relative to ABI's current capabilities (Lindsey et al. 2023). Of these upgrades, an infrared window channel with an increased 1-km nadir pixel size and a visible channel with 0.25-km nadir pixel size (although with required MTFs corresponding to 1.5 and 0.3 km, respectively) are included. These changes have the potential to improve SR capabilities due to the current lack of higher-resolution channels at infrared wavelengths on ABI. However, these changes may also pose challenges in reconciling an 8-times scale difference between the channels with the highest (0.25 km) and lowest (2.0 km) pixel sizes.

6. Conclusions

Our results indicate that we are able to superresolve all channels on ABI to a common 0.5-km resolution and substantially outperform bicubic interpolation as a baseline. The presented approach uses the native 0.5-km visible channel on ABI to guide the superresolution of the 1- and 2-km channels. The amount of value added by the CNN depends heavily on the channel. For example, when evaluated on a reduced-resolution dataset, we observe a roughly 17.9-times-lower RMSE for band 1 ($0.47\ \mu\text{m}$), a roughly 5.6-times-lower RMSE for band 6 ($2.20\ \mu\text{m}$), and the smallest value-add of 2.5-times-lower RMSE for band 8 ($6.20\ \mu\text{m}$). Comparisons using structural similarity index measure show large improvements in overall image quality. We also show that the CNN produces radiance spectra that are overall more similar to reference imagery compared to bicubic interpolation. Comparisons with high-resolution imagery from *Landsat-8* and *Landsat-9* OLI/TIRS instruments show overall good correspondence with the superresolved imagery. However, there is only minor improvement in ABI band 4 and some issues with atmospheric features that ABI band 2 has little sensitivity to. Estimating all ABI bands at 0.5-km resolution may allow for more easily combining information across bands without reconciling differences in spatial resolution. However, more analysis is needed to determine potential impacts on derived products or multispectral imagery that use superresolved imagery. This approach is extensible to other remote sensing instruments and requires only a small amount of data and knowledge of each channel's spatial response.

Acknowledgments. This work was supported by the NOAA GOES-R and GeoXO Programs under Grant NA19OAR4320073. C.H.W. and I.E.-U. also acknowledge the support by the National Science Foundation under Grant ICER-2019758.

Data availability statement. The GOES-16 ABI data used in this work can be accessed from <https://registry.opendata.aws/noaa-goes>. Landsat data can be accessed through the EarthExplorer interface at <https://earthexplorer.usgs.gov/>.

REFERENCES

- Armansson, S. E., M. O. Ulfarsson, J. Sigurdsson, H. V. Nguyen, and J. R. Sveinsson, 2021: A comparison of optimized Sentinel-2 super-resolution methods using Wald's protocol and Bayesian optimization. *Remote Sens.*, **13**, 2192, <https://doi.org/10.3390/rs13112192>.
- Barsi, J. A., K. Lee, G. Kvaran, B. L. Markham, and J. A. Pedelty, 2014: The spectral response of the Landsat-8 operational land imager. *Remote Sens.*, **6**, 10232–10251, <https://doi.org/10.3390/rs61010232>.
- Boreman, G. D., 2001: *Modulation Transfer Function in Optical and Electro-Optical Systems*. Vol. 4, SPIE Press, 110 pp.
- Calle, A., J.-L. Casanova, and F. González-Alonso, 2009: Impact of point spread function of MSG-SEVIRI on active fire detection. *Int. J. Remote Sens.*, **30**, 4567–4579, <https://doi.org/10.1080/01431160802609726>.
- Deborah, H., N. Richard, and J. Y. Hardeberg, 2015: A comprehensive evaluation of spectral distance functions and metrics for hyperspectral image processing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **8**, 3224–3234, <https://doi.org/10.1109/JSTARS.2015.2403257>.
- Deneke, H., and Coauthors, 2021: Increasing the spatial resolution of cloud property retrievals from Meteosat SEVIRI by use of its high-resolution visible channel: Implementation and examples. *Atmos. Meas. Tech.*, **14**, 5107–5126, <https://doi.org/10.5194/amt-14-5107-2021>.
- Dong, C., C. C. Loy, K. He, and X. Tang, 2015: Image super-resolution using deep convolutional networks. arXiv, 1501.00092v3, <https://doi.org/10.48550/arXiv.1501.00092>.
- Drusch, M., and Coauthors, 2012: Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.*, **120**, 25–36, <https://doi.org/10.1016/j.rse.2011.11.026>.
- Farsiu, S., M. D. Robinson, M. Elad, and P. Milanfar, 2004: Fast and robust multiframe super resolution. *IEEE Trans. Image Process.*, **13**, 1327–1344, <https://doi.org/10.1109/TIP.2004.834669>.
- Freeman, W. T., T. R. Jones, and E. C. Pasztor, 2002: Example-based super-resolution. *IEEE Comput. Graph. Appl.*, **22**, 56–65, <https://doi.org/10.1109/38.988747>.
- Gargiulo, M., A. Mazza, R. Gaetano, G. Ruello, and G. Scarpa, 2019: Fast super-resolution of 20 m Sentinel-2 bands using convolutional neural networks. *Remote Sens.*, **11**, 2635, <https://doi.org/10.3390/rs11222635>.
- Geiss, A., and J. C. Hardin, 2020: Radar super resolution using a deep convolutional neural network. *J. Atmos. Oceanic Technol.*, **37**, 2197–2207, <https://doi.org/10.1175/JTECH-D-20-0074.1>.
- , and —, 2023: Strictly enforcing invertibility and conservation in CNN-based super resolution for scientific datasets. *Artif. Intell. Earth Syst.*, **2**, e210012, <https://doi.org/10.1175/AIES-D-21-0012.1>.
- Glasner, D., S. Bagon, and M. Irani, 2009: Super-resolution from a single image. *2009 IEEE 12th Int. Conf. on Computer Vision*, Kyoto, Japan, Institute of Electrical and Electronics Engineers, 349–356, <https://doi.org/10.1109/ICCV.2009.5459271>.
- Green, R. O., and Coauthors, 1998: Imaging spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sens. Environ.*, **65**, 227–248, [https://doi.org/10.1016/S0034-4257\(98\)00064-9](https://doi.org/10.1016/S0034-4257(98)00064-9).
- He, K., X. Zhang, S. Ren, and J. Sun, 2015: Deep residual learning for image recognition. arXiv, 1512.03385v1, <https://doi.org/10.48550/arXiv.1512.03385>.
- He, L., J. Zhu, J. Li, D. Meng, J. Chanussot, and A. Plaza, 2020: Spectral-fidelity convolutional neural networks for hyperspectral pansharpening. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **13**, 5898–5914, <https://doi.org/10.1109/JSTARS.2020.3025040>.
- Hillger, D., and Coauthors, 2013: First-light imagery from Suomi NPP VIIRS. *Bull. Amer. Meteor. Soc.*, **94**, 1019–1029, <https://doi.org/10.1175/BAMS-D-12-00097.1>.
- Huang, H., and Coauthors, 2020: UNet 3+: A full-scale connected UNet for medical image segmentation. arXiv, 2004.08790v1, <https://doi.org/10.48550/arXiv.2004.08790>.
- Ioffe, S., and C. Szegedy, 2015: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv, 1502.03167v3, <https://doi.org/10.48550/arXiv.1502.03167>.
- Irons, J. R., J. L. Dwyer, and J. A. Barsi, 2012: The next Landsat satellite: The Landsat Data Continuity Mission. *Remote Sens. Environ.*, **122**, 11–21, <https://doi.org/10.1016/j.rse.2011.08.026>.
- Javan, F. D., F. Samadzadegan, S. Mehravar, A. Toosi, R. Khatami, and A. Stein, 2021: A review of image fusion techniques for pansharpening of high-resolution satellite imagery. *ISPRS J. Photogramm. Remote Sens.*, **171**, 101–117, <https://doi.org/10.1016/j.isprsjprs.2020.11.001>.
- Kruse, F. A., A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz, 1993: The Spectral Image Processing System (SIPS)—Interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ.*, **44**, 145–163, [https://doi.org/10.1016/0034-4257\(93\)90013-N](https://doi.org/10.1016/0034-4257(93)90013-N).
- Lanaras, C., J. Bioucas-Dias, S. Galliani, E. Baltsavias, and K. Schindler, 2018: Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network. *ISPRS J. Photogramm. Remote Sens.*, **146**, 305–319, <https://doi.org/10.1016/j.isprsjprs.2018.09.018>.
- Ledig, C., and Coauthors, 2017: Photo-realistic single image super-resolution using a generative adversarial network. arXiv, 1609.04802v5, <https://doi.org/10.48550/arXiv.1609.04802>.
- Li, H., Y. Yang, M. Chang, S. Chen, H. Feng, Z. Xu, Q. Li, and Y. Chen, 2022: SRDIFF: Single image super-resolution with diffusion probabilistic models. *Neurocomputing*, **479**, 47–59, <https://doi.org/10.1016/j.neucom.2022.01.029>.
- Lim, B., S. Son, H. Kim, S. Nah, and K. M. Lee, 2017: Enhanced deep residual networks for single image super-resolution. arXiv, 1707.02921v1, <https://doi.org/10.48550/arXiv.1707.02921>.
- Lindsey, D., and Coauthors, 2023: GXI: NOAA's Geostationary Imager of the Future. *EUMETSAT Satellite Conf.*, Malmö, Sweden, GeoXO, <https://www-cdn.eumetsat.int/files/2023-10/5.%20Dan%20Lindsey%2015.00.pdf>.
- Liu, H., Z. Ruan, P. Zhao, C. Dong, F. Shang, Y. Liu, L. Yang, and R. Timofte, 2022: Video super resolution based on deep learning: A comprehensive survey. arXiv, 2007.12928v3, <https://doi.org/10.48550/arXiv.2007.12928>.
- Liu, Q., L. Han, R. Tan, H. Fan, W. Li, H. Zhu, B. Du, and S. Liu, 2021: Hybrid attention based residual network for pansharpening. *Remote Sens.*, **13**, 1962, <https://doi.org/10.3390/rs13101962>.
- Lu, Z., and Y. Chen, 2019: Single image super resolution based on a modified U-net with mixed gradient loss. arXiv, 1911.09428v1, <https://doi.org/10.48550/arXiv.1911.09428>.

- Masi, G., D. Cozzolino, L. Verdoliva, and G. Scarpa, 2016: Pansharpening by convolutional neural networks. *Remote Sens.*, **8**, 594, <https://doi.org/10.3390/rs8070594>.
- Ronneberger, O., P. Fischer, and T. Brox, 2015: U-Net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, Lecture Notes in Computer Science, Vol. 9351, Springer, 234–241, https://doi.org/10.1007/978-3-319-24574-4_28.
- Salgueiro Romero, L., J. Marcello, and V. Vilaplana, 2020: Super-resolution of Sentinel-2 imagery using generative adversarial networks. *Remote Sens.*, **12**, 2424, <https://doi.org/10.3390/rs12152424>.
- Scarpa, G., S. Vitale, and D. Cozzolino, 2018: Target-adaptive CNN-based pansharpening. *IEEE Trans. Geosci. Remote Sens.*, **56**, 5443–5457, <https://doi.org/10.1109/TGRS.2018.2817393>.
- Schmit, T. J., and M. M. Gunshor, 2021: GOES-R Advanced Baseline Imager (ABI) algorithm theoretical basis document for Cloud and Moisture Imagery Product (CMIP). NOAA NESDIS, Center for Satellite Applications and Research Doc., version 4.0, 58 pp., https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Enterprise/ATBD_Enterprise_Cloud_and_Moisture_Imagery_Product_v4_2021-01-13.pdf.
- , P. Griffith, M. M. Gunshor, J. M. Daniels, S. J. Goodman, and W. J. Lebar, 2017: A closer look at the ABI on the GOES-R series. *Bull. Amer. Meteor. Soc.*, **98**, 681–698, <https://doi.org/10.1175/BAMS-D-15-00230.1>.
- Stengel, K., A. Glaws, D. Hettinger, and R. N. King, 2020: Adversarial super-resolution of climatological wind and solar data. *Proc. Natl. Acad. Sci. USA*, **117**, 16 805–16 815, <https://doi.org/10.1073/pnas.1918964117>.
- Szegedy, C., S. Ioffe, V. Vanhoucke, and A. Alemi, 2016: Inception-v4, inception-ResNet and the impact of residual connections on learning. arXiv, 1602.07261v2, <https://doi.org/10.48550/arXiv.1602.07261>.
- Tsakatakis, G., A. Aidini, K. Fotiadou, M. Giannopoulos, A. Pentari, and P. Tsakalides, 2019: Survey of deep-learning approaches for remote sensing observation enhancement. *Sensors*, **19**, 3929, <https://doi.org/10.3390/s19183929>.
- Vandal, T., E. Kodra, S. Ganguly, A. Michaelis, R. Nemani, and A. R. Ganguly, 2017: DeepSD: Generating high resolution climate change projections through single image super-resolution. arXiv, 1703.03126v1, <https://doi.org/10.48550/arXiv.1703.03126>.
- Wald, L., 2002: *Data Fusion: Definitions and Architectures—Fusion of Images of Different Spatial Resolutions*. Presses de l'Ecole, Ecole des Mines de Paris, 200 pp.
- Wang, D., Y. Li, L. Ma, Z. Bai, and J. C.-W. Chan, 2019: Going deeper with densely connected convolutional neural networks for multispectral pansharpening. *Remote Sens.*, **11**, 2608, <https://doi.org/10.3390/rs11222608>.
- Wang, Z., A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, 2004: Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.*, **13**, 600–612, <https://doi.org/10.1109/TIP.2003.819861>.
- Wu, J., Z. He, and J. Hu, 2020: Sentinel-2 sharpening via parallel residual network. *Remote Sens.*, **12**, 279, <https://doi.org/10.3390/rs12020279>.
- , L. Lin, C. Zhang, T. Li, X. Cheng, and F. Nan, 2023: Generating Sentinel-2 all-band 10-m data by sharpening 20/60-m bands: A hierarchical fusion network. *ISPRS J. Photogramm. Remote Sens.*, **196**, 16–31, <https://doi.org/10.1016/j.isprsjprs.2022.12.017>.
- Wu, X., and T. Schmit, 2019: GOES-16 ABI Level 1b and Cloud and Moisture Imagery (CMI) release full validation data quality. NOAA Tech. Doc., 47 pp., https://www.ncei.noaa.gov/sites/default/files/2021-08/GOES-16_ABI-L1b-CMI_Full-Validation_ProductPerformanceGuide_v2.pdf.
- Wulder, M. A., and Coauthors, 2019: Current status of Landsat program, science, and applications. *Remote Sens. Environ.*, **225**, 127–147, <https://doi.org/10.1016/j.rse.2019.02.015>.
- Yang, W., X. Zhang, Y. Tian, W. Wang, J.-H. Xue, and Q. Liao, 2019: Deep learning for single image super-resolution: A brief review. *IEEE Trans. Multimedia*, **21**, 3106–3121, <https://doi.org/10.1109/TMM.2019.2919431>.
- Zhou, Z., M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, 2018: UNet++: A nested U-net architecture for medical image segmentation. arXiv, 1807.10165v1, <https://doi.org/10.48550/arXiv.1807.10165>.