Lightweight Super-Resolution of Satellite Imagery for Nitrogen Dioxide Prediction

Zachary Yahn

University College Dublin, Dublin, Ireland

Abstract. Nitrogen dioxide is a common air pollutant that can be difficult to track with traditional ground-based monitoring. While satellites can provide plentiful data and global coverage, such images are often captured at low resolutions that prohibit fine-grained predictions. In this work we investigate the application of super-resolution to the nitrogen dioxide prediction problem by super-resolving Sentinel-2 images with several popular models and our own lightweight architecture. Instead of collecting a second high-resolution dataset for ground-truths, we explore two methods for super-resolving images with a single dataset, thus assessing the viability of a greatly simplified super-resolution pipeline. Furthermore, we investigate the relationship between common super-resolution metrics such as per-pixel error and perceptual quality with empirical performance on the downstream nitrogen dioxide prediction task. We find that our bespoke very small model, NinaCustom, outperforms popular deep models on common error benchmarks while offering inference time speedups of up to 50x. We observe that our approach has the potential to slot into the pre-processing steps of numerous computer vision tasks without adding a significant amount of computational overhead or requiring additional data collection while increasing downstream performance.

Keywords: Super-resolution · Nitrogen dioxide · Satellite imagery.

1 Introduction

Air pollution is a pressing public health concern that is becoming increasingly problematic given continued industrialization, urbanization, and population growth around the world. The United States Environmental Protection Agency (EPA) recognises several prevalent pollutant compounds, including particulate matter ($PM_{2.5}$), carbon monoxide (CO), and nitrogen dioxide (NO_2) [4]. Khajeamiri et al., observe that NO_2 has a positive correlation with mortality, severity, and transmission of various viral respiratory infections [16]. They also find that it increases the risk of asthma, and that NO_2 can have a stronger and faster effect on pneumonia and bronchitis than other air pollutants. Children exposed to NO_2 are at higher risk for respiratory viral infections as well [16]. The Global Bureau of Disease recognizes air pollution (both ambient and household) as a severe health risk associated with premature mortality [19].

Despite its apparent effects on public health, understanding and anticipating NO_2 concentration is challenging. Lamsal et al. show via satellite imagery that NO_2 density scales as a power law with population size in urban areas. However, they find that population density alone is insufficient for predicting nitrogen dioxide [17]. In a case study of Ulaanbaatar, the urban capital of Mongolia, researchers show that land use variables such as proximity to the city center, density of roads, and presence of power plants can also predict NO_2 concentration. Furthermore, they show that NO_2 levels are highly dependent on the season [15]. Arain et al. show that road networks are a major cause of increased NO_2 concentration, while sensors as far as 300m from a major highway did not detect an increase in NO_2 above nominal atmospheric levels even in the middle of the urban area chosen for testing [3]. To summarize, accurately detecting nitrogen dioxide over large areas requires fine-grained sensor data. While government agencies like the EPA and European Environmental Agency (EEA) have established monitoring stations capable of detecting NO_2 content, these lack the resolution and coverage to track nitrogen dioxide on national or global scales.

One solution to this problem is the use of satellite imagery, which has the potential to cover a much broader area than monitoring stations. While using satellites addresses coverage, high-resolution satellite data are expensive to obtain. Low-resolution datasets, on the other hand, are plentiful. Super-resolution, a deep learning technique that uses artificial neural networks to intelligently upscale images, may be able to enhance low-resolution satellite data to make them more useful for NO_2 prediction. Super resolution can be broadly divided into two approaches: single image super-resolution (SISR) and multi-image super-resolution (MISR). In SISR, a single image is upscaled by training a deep neural network to interpolate values or fill in pixels with latent feature information. MISR, on the other hand, combines information from multiple images, for example the same scene at different times, to synthesize a higher resolution output. In either case, a deep learning model learns to map the input low resolution image(s) into a high resolution image using labeled examples, either from another data source or by downscaling the given data and using the originals as ground-truths [25]. Super-resolution techniques often require an additional high-resolution dataset of the same scenes to use as ground-truth labels, thereby limiting the broad range of low-resolution images to a narrower subset that can be matched to another satellite.

In this study we show that applying super-resolution to satellite images and then providing those images to a nitrogen dioxide prediction model yields improved performance over using the unaltered images, and we conduct experiments to investigate the relationships between model size and architecture with this improvement. We also explore two ways to super-resolve such images without the use of an additional dataset. All of our code is available in our Github repository. [1]

[1] https://github.com/zacharyyahn/Nitrogen-SR

1.1 Related Works

Several studies have investigated the use of satellite imagery from various sources for predicting air pollution. In their work, Sorek et al. use a modified VGG-16 network to create air quality maps for NO_2 and other pollutants using WorldView-2 [10] imagery [27]. Similarly, Zhu et al. develop a deep random forest for predicting nitrogen dioxide levels in China [30]. They use data from the tropospheric monitoring instrument (TROPOMI) onboard Sentinel-5p [9]. The most popular data source for air pollution prediction appears to be Sentinel-2 [8]. Scheibenreif et al. use a combination of Sentinel-2 and Sentinel-5p data and a modified ResNet50 model to predict NO_2 concentration in Western Europe [23]. Likewise, Rowley et al. use the same dataset with the addition of tabular weather and seasonal data to improve upon their approach [22].

Other studies instead explore applications of super-resolution to satellite imagery. In their study, Liu et al., use SISR to synthesize data from mobile, ground-based air pollution sensors attached to vehicles in a city [20]. While these data do not come from a satellite, the application of super-resolution to this task demonstrates that the technique may be applicable to air pollution. Galar et al. develop a model called an enhanced deep residual network (EDSR) that avails of residual connections to allow for a deeper network. They use this model for SISR on Sentinel-2 imagery, with RapidEye [6] data for ground-truths [11]. Lanaras et al. investigate traditional convolutional neural networks (CNNs) individualized to channel resolutions. Instead of a ground-truth set from another source, they use high-resolution channels as labels for low-resolution channels in the same image [18]. Valsesia et al. deploy an MISR model that achieves invariance to temporal ordering in input images, recording a top score on the European Space Agency's (ESA) Proba-V [5] super-resolution challenge. They also introduce a method for characterising the uncertainty of their model by adding an additional output unit that predicts bias [28]. Other studies explore alternatives to CNN-based super-resolution. For example, Arefin et al. use a gated recurrent unit (GRU) with a convolutional input layer that also achieves a top score on Proba-V [2]. Transformer models have also been shown to be effective on the same dataset [1].

Finally, a few studies show that super-resolved images may also increase performance in downstream tasks. Shermeyer et al. use CNN-based SISR to super-resolve images, and then show that an object detection model trained on these images achieves a stronger performance [26]. Similarly, Razzak et al. use an MISR model to enhance Sentinel-2 images, and then observe that these images improve the accuracy of a semantic segmentation model trained for building delineation [21].

1.2 Contributions of the Paper

If super-resolution can be simplified to not require additional data and use a lightweight model, then it may be an effective pre-processing step in the pipeline for countless satellite imagery tasks. In this work, we first investigate whether

super-resolved images can improve performance on the important task of nitrogen dioxide prediction. While several studies have successfully predicted nitrogen dioxide pollution from satellite imagery, and multiple others have shown that super-resolution techniques can improve satellite images, it remains unclear if these combined approaches would yield improved results. It is also unclear which super-resolution models would be effective and how large they need to be; the use of very large models may reduce the viability of super-resolution as a common pre-processing steps for such tasks due to computational intensity. Furthermore, reliance on additional datasets limits the scope of super-resolution model training and increases the complexity of data collection. Specifically, we identify three key shortcomings that to our knowledge have not been addressed in prior literature:

1. Overemphasis on statistical image quality metrics instead of empirical performance on downstream tasks.
2. Reliance on additional higher-resolution datasets for ground-truths.
3. Dependence on large models without testing lightweight versions.

To investigate these shortcomings, we train several super-resolution models on a single dataset, using high-resolution channels as labels for low-resolution ones and eliminating the need for an additional dataset. We test the super-resolved images on a nitrogen dioxide prediction task to asses super-resolution model performance. Finally, we develop our own custom ultra-lightweight super-resolution model to assess the effect of a wide range of parameter counts on performance, and we compare it with several state-of-the-art architectures.

Fig. 1. Locations of air pollution monitoring stations that provide ground-truth nitrogen dioxide information.

2 Dataset

To standarize comparison between our methods and other nitrogen dioxide prediction tasks, we use the dataset curated by Scheibenreif et al. [24]. The set contains Sentinel-2 frames that are temporally aligned with observations from air pollution monitoring stations on monthly, quarterly, and annual bases. These stations are managed by the United States Environmental Protection Agency (EPA) and are in various locations on the West Coast of the United States. Figure 1 shows a map of these locations.

The Sentinel-2 mission has two identical satellites, Sentinel-2A and Sentinel-2B, which have been collecting images of Earth's surface since June, 2015. The satellites fly in the same orbit but out of phase by 180°, giving a high equatorial revisit frequency of five days. Each satellite carries optical instruments that collect thirteen channels: four bands at 10m, six bands at 20m, and three bands at 60m. Scheibenreif et al. do not include Band 10, one of the 60m bands, in their dataset, which leaves twelve channels total per image [24]. Figure 2 shows an example of these twelve bands for a single location, including the wavelength of each band [7].

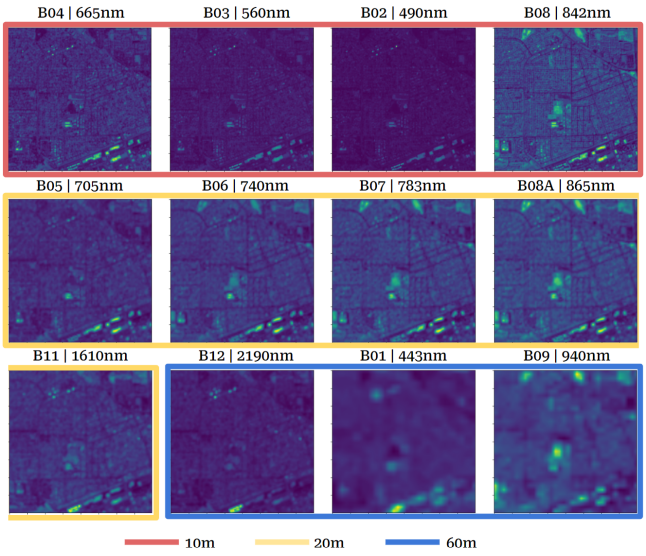

Fig. 2. Example of twelve Sentinel-2 bands for one location.

In their preparation of the dataset, Scheibenreif et al. use bicubic interpolation to upscale the 20m and 60m bands to the same resolution as the 10m bands, yielding a final 12×200×200 image for each ground observation measurement. There are a total of 3,276 ground station observations for the 2018-2020 period recorded in the dataset and thus the same number of images.

3 Methods

3.1 Overview of Proposed Framework

Our framework is composed of two parts: super-resolution and nitrogen dioxide prediction. First, a dataset of super-resolved images are created using one of two methods, described below. Then, these images are provided to a nitrogen prediction model that uses the ground station observations as ground-truths. Figure 3 contains a diagram that outlines our framework.

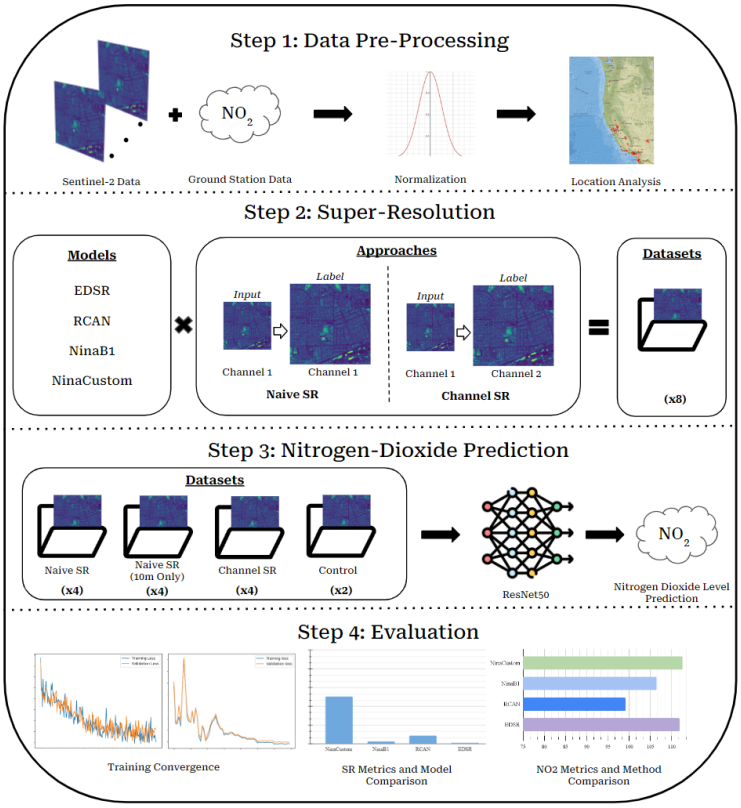

Fig. 3. Diagram of proposed framework.

3.2 Super-Resolution

We super-resolve the dataset using four different models. Two of these, the Enhanced Deep Residual Network (EDSR) [11] and Residual Channel Attention Network (RCAN) [29] are state of the art super-resolution models. EDSR is the

largest, using residual connections to backpropagate the error through a deeper model. RCAN uses a residual in residual (RIR) structure with attention layers to learn relationships between representations of low-frequency information. We also use the NinaB1 version of NinaSR, a family of small super-resolution models that make use of a combination of convolutional and attentional layers [12]. All three of these models are implemented in the torchsr library [13]. Finally, we develop our own custom Nina model by adapting the code for NinaB1 to reduce the number of representational features and the number of attention layers. The result is an ultra-lightweight model that offers significantly faster training and inference time over the other versions. All four models, their sizes, and their advantages are summarized below in Table 1.

Table 1. Four models used in this study.

Model	No. Parameters	Advantages
EDSR	40,729,627	Very large model capable of deep feature extraction.
RCAN	15,444,667	Large model extensively leveraging attention.
NinaB1	1,019,794	Balanced model that trades-off speed and depth.
NinaCustom	51,178	Small model prioritizing speed.

In addition to using four different models, we super-resolve the Sentinel-2 images according to two different methodologies that we dub naive super-resolution and channel super-resolution. Between three methods and four models for each method we thus conduct eight super-resolution experiments. In all cases, all input images are normalized according to:

$$x_{i,c} = \frac{x_{i,c} - \mu_c}{\sigma_c} \tag{1}$$

Where $x_{i,c}$ is image i for channel c, μ_c is the average value for that channel, and σ_c is the standard deviation of that channel. For all experiments the data are divided into training, validation, and test sets corresponding to 60%, 20%, and 20% of the full dataset, respectively. All models with all approaches are trained for 200 epochs with a batch size of one, and a learning rate of 0.0001 with a mean-squared error (MSE) loss function and the Adam optimizer.

Naive Super-Resolution. For naive super-resolution (Naive SR), we train the same model on all twelve channels indiscriminately. We theorize that the model will learn to improve low- and high-resolution channels alike by being exposed to all twelve during training. Each channel of each image is provided separately as a two-dimensional 200×200 array, giving 39,312 image from the 3,276 observations. The input to the model is a downscaled version of each image, and the label is the image at its original resolution. Once the model is trained, it is then applied to the original images to produce a new dataset of super-resolved images of size 12×400×400.

Channel Super-Resolution. Unlike naive super-resolution, channel super-resolution (Channel SR) uses high-resolution 10m channels as labels for the lower-resolution 20m channels. The inputs to the model are 12×100×100 images that have been downsampled to their original size (undoing the upsampling performed during dataset curation by Scheibenreif et al.). The model then learns to increase the resolution of each 20m channel using the 10m channel that is closest to it in terms of observational wavelength, similar to the approach used in [18]. This gives a final set of images of size 10×200×200 for each observation in the dataset. The 60m channels are discarded for this approach because they do not have analagous 10m channels of similar wavelengths.

3.3 Super-Resolution Evaluation Metrics

Each model and each approach is evaluated with three metrics. Mean-squared error (MSE) is used to evaluate performance on a per-pixel level. Peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) are used to measure overall perceptual quality of the images. These metrics correspond to the following equations:

$$MSE(x_i, y_i) = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2 \tag{2}$$

$$PSNR(x_i, y_i) = 10\log_{10}\frac{255}{MSE(x_i, y_i)} \tag{3}$$

$$SSIM(x_i, y_i) = \frac{(2\mu_x\mu_y + 0.01)(\sigma_{xy} + 0.03)}{(\mu_x^2 + \mu_y^2 + 0.01)(\sigma_x^2 + \sigma_y^2 + 0.03)} \tag{4}$$

Where x_i and y_i are the predicted and label images, respectively, n is the number of images in the set, μ_x and μ_y are the means of a given predicted or label image, σ_x and σ_y are the standard deviations of a given predicted or label image, and σ_{xy} is the covariance of the predicted and label images.

3.4 Nitrogen Dioxide Prediction

The super-resolution stage of our framework produces eight datasets corresponding to super-resolved images for every combination of model and approach. In addition to prediction experiments on these eight datasets, we also conduct four more experiments using only the 10m channels from the naive super-resolution outputs. Since it is unclear from prior work which channels are most important for nitrogen dioxide prediction, we theorize that the highest-resolution channels will contain the most information. Finally, we use two control datasets that use images at their original resolution: one of the original images, and one of just the 10m channels of the original images. Thus in total we have 14 datasets. Each dataset was then provided to a ResNet50 [14] model with its output layer modified to predict a single value. As with super-resolution, each dataset was divided into training, validation, and testing sets with a 60-20-20 split. Input images

and nitrogen dioxide labels were both normalized according to (1). Initial experimentation showed that training performance was very susceptible to learning rate, therefore a grid search over learning rates was performed for each dataset. Learning rates 0.00005, 0.000075, 0.0001, 0.00025, 0.0005, and 0.001 were tested, each for 70 epochs with a batch size of 1, the Adam optimizer, and a learning rate that decayed by 0.5 every 10 epochs. A mean absolute error (MAE) loss function was used according to the following equation:

$$MAE(x_i, y_i) = \frac{1}{n} \sum_{i=1}^{n} |x_i - y_i| \tag{5}$$

3.5 Nitrogen Dioxide Prediction Evaluation Metrics

Nitrogen dioxide levels predicted by the ResNet50 were compared to the ground-truth values observed by ground monitoring stations. These were then rescaled and evaluated using MSE (2) and MAE (5).

4 Results

4.1 Super-resolution

Figure 4 shows the loss on the training and validation sets during training. Although all models eventually converged, those performing super-resolution via ChannelSR struggled much more, with the notable exception of NinaCustom. The figure shows that all models successfully learned to upscale Sentinel-2 images without overfitting to their training data.

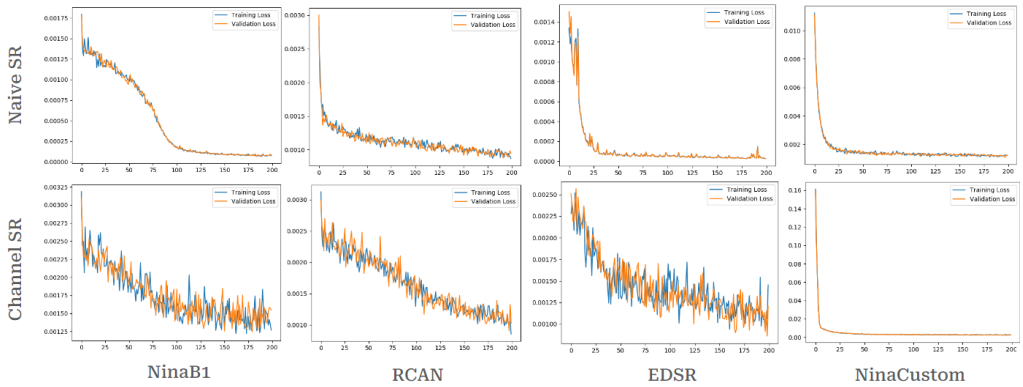

Fig. 4. Training loss and validation loss for each combination of super-resolution model and approach. Training loss is in blue, and validation loss is in orange.

The results of scoring all models and approaches on MSE, PSNR, and SSIM are shown in Figure 5. For Naive super-resolution, NinaCustom scores a noticeably higher MSE, but is more comparable when it comes to SSIM and PSNR. However, NinaB1, RCAN, and EDSR all achieve an SSIM close to 1.0, while NinaCustom achieves 0.72. For the channel super-resolution approach NinaCustom once again achieves the worst results, but in this case is more comparable with the other models. Notably, all four score nearly the same PSNR. Since the channel super-resolution approach uses high-resolution channels as labels for low-resolution ones, achieving a near-perfect SSIM is not necessarily desirable; this might constitute losing the information that is unique to the low-resolution channels. Instead, we theorize that this approach much strike a balance between increasing the resolution of the images without overwriting their unique features. While these metrics provide useful insight into the performance of various super-resolution models, for our purposes it is more important how these images influence performance on the downstream nitrogen-dioxide prediction task.

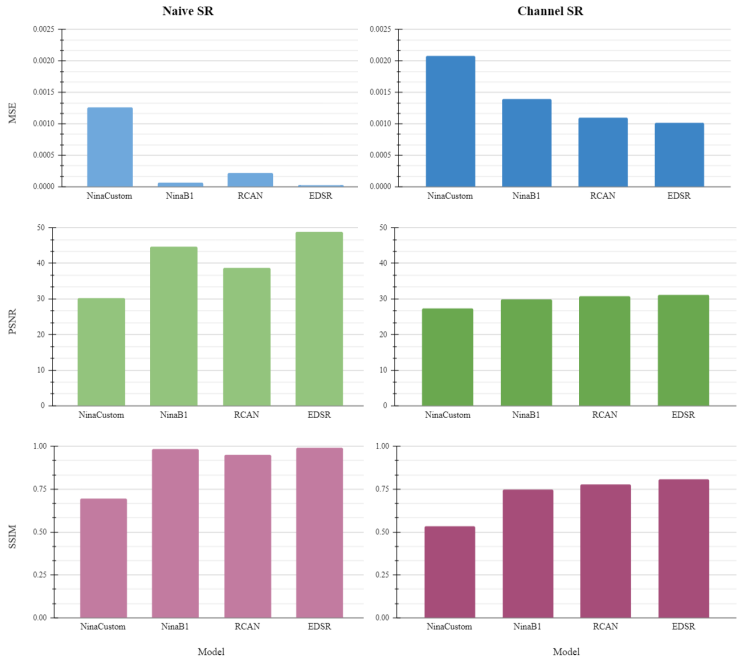

Fig. 5. MSE, PSNR, and SSIM performance on held-out test set for each super-resolution method and model.

4.2 Nitrogen Dioxide Prediction

Figure 6 shows the training and validation loss curves during training for the best hyperparameters found during the grid search. Although some of the models struggled to learn at first, all eventually managed to converge.

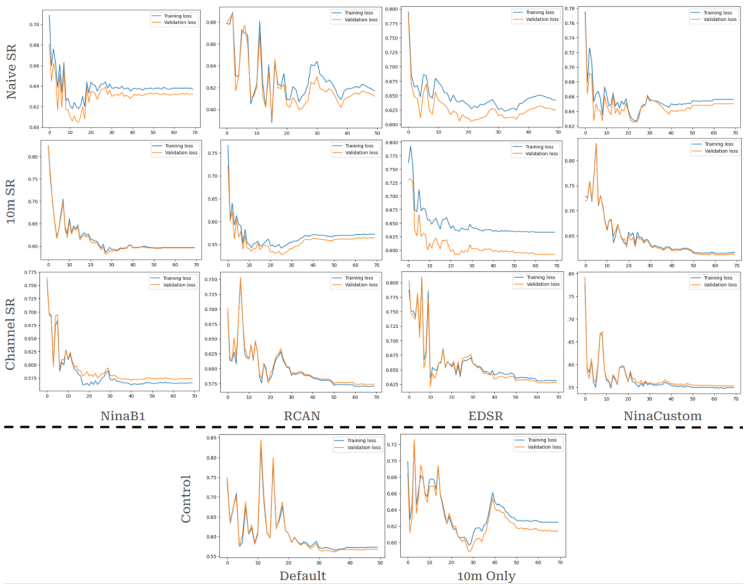

Fig. 6. Training loss and validation loss for a ResNet50 model trained on each super-resolved dataset. Training loss is in blue, and validation loss is in orange.

Finally, we collect the MSE and MAE scores for all nitrogen prediction models in Figure 7. Across all super-resolution approaches RCAN achieves the lowest MSE and MAE. Notably, using only the 10m channels and applying RCAN for super-resolution had the best results out of any combination of model and approach. With the exception of the experiments involving EDSR, the channel super-resolution approach had the best results across all models. It was also the only approach where NinaCustom achieved lower MSE and MAE scores than control experiments. NinaCustom also outperformed the other three models on the channel super-resolution approach, achieving the second lowest MSE and MAE across all combinations.

5 Discussion

The results in Figure 5 show that the super-resolution models perform as one would expect. On naive super-resolution the largest model, EDSR, achieves lower MSE and higher PSNR and SSIM scores than all other models. NinaCustom

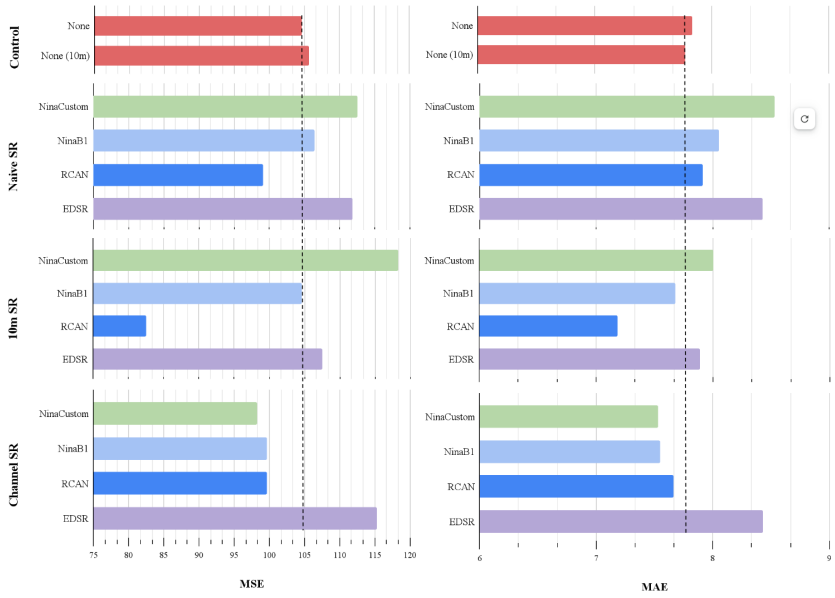

Fig. 7. Results of nitrogen dioxide prediction for all models. Note axis bounds are adjusted for both MSE and MAE. Vertical dashed lines added to compare to most effective control result.

scores especially poorly across all metrics, which is expected given that it is 5% of the next largest model. This seems to indicate that, at least for a classic super-resolution scenario, a larger model achieves better scores on the typical evaluation metrics.

When it comes to channel super-resolution the results are similar: EDSR outscores the other models, with NinaCustom achieving the poorest results. Notably, across all models and metrics the channel super-resolution approach achieves worse scores than naive super-resolution. While there is a more significant range of performances with the pixel-wise MSE metric, both perceptual metrics seem so show that the model architecture has less effect on performance. Especially with PSNR, there seems to be a limit to how well a model can learn to map low-resolution channels to high-resolution channels, indicating that larger models may not be necessary for this approach.

It is also important to consider how efficiently each model can super-resolve images. Figure 8 shows the amount of time each model requires to upscale 500 200×200 images when run on an Intel i7 CPU. Note that NinaCustom offers significant speedups over the larger models. This speed increase may be beneficial in scenarios where large amounts of data must be processed, for example predicting nitrogen dioxide levels on a national or global scale. It may also be useful for researchers with limited computational resources at their disposal. While the faster NinaCustom model does achieve worse scores, particularly in naive super-

resolution, the question remains whether these scores translate to performance on the nitrogen dioxide prediction task.

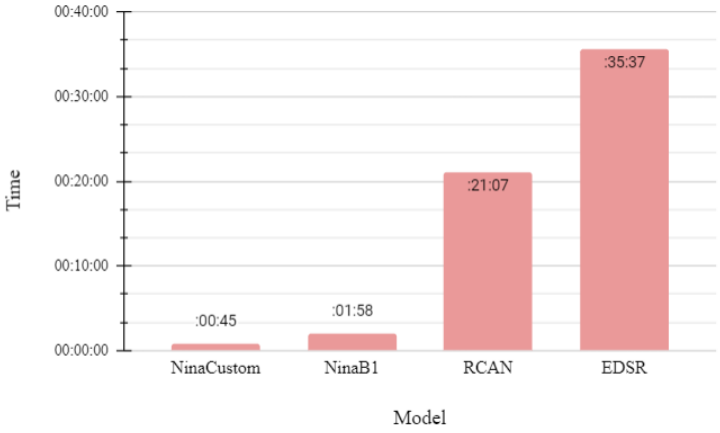

Fig. 8. Inference time for 500 images for each model

One can see from the results in Figure 7 that the best performing super-resolution models do not necessarily produce the best images for nitrogen dioxide prediction. Notably, the images upscaled by EDSR models achieve the lowest MSE and MAE scores across nearly all methods. RCAN produces the best results out of any combination despite having lower super-resolution scores than NinaB1 or EDSR. The results also show that NinaCustom outperforms all other models on the channel super-resolution approach despite achieving a much worse performance across all super-resolution methods. While it seems that larger models and specific architectures like RCAN are more appropriate for a naive super-resolution task, using high-resolution channels to super-resolve low-resolution channels can avail of smaller models and achieve comparable results to the naive approach. We can conclude that a very small super-resolution model like NinaCustom can effectively upscale images and increase performance on a downstream nitrogen prediction task with minimal computation and temporal cost.

6 Conclusion

In this study we compare several super-resolution models with our own custom lightweight model across two approaches to upscaling Sentinel-2 imagery for nitrogen dioxide prediction. We then use the super-resolved images as inputs to a ResNet50 model, finding that RCAN is the most effective model for super-resolving all channels by learning to increase the resolution of a downscaled version to match the original image, while our small NinaCustom model achieves the best results when using high-resolution channels as labels for low-resolution ones.

Notably, this approach, combined with the NinaCustom models, also achieves significant speedups over all other models in our study without sacrificing performance. This makes it ideal in situations with reduced computational resources or for processing large datasets.

While this study compares several models combined with several approaches to super-resolution, future studies may also include other popular architectures like transformers or diffusion models. Future work might also investigate alternative lightweight architectures to NinaCustom. We invite other researchers exploring nitrogen dioxide prediction and potentially similar tasks as well to consider employing small super-resolution models to enhance the quality of satellite imagery datasets.

References

1. An, T., Zhang, X., Huo, C., Xue, B., Lingfeng: Tr-misr: Multiimage super-resolution based on feature fusion with transformers. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **15**, 1373–1388 (2022). https://doi.org/10.1109/JSTARS.2022.3143532
2. Arefin, M.R., Michalski, V., St-Charles, P.L., Kalaitzis, A., Kim, S., Kahou, S.E., Bengio, Y.: Multi-image super-resolution for remote sensing using deep recurrent networks. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops pp. 816–825 (2020). https://doi.org/10.1109/CVPRW50498.2020.00111
3. Beckerman, B., Jerrett, M., Brook, J.R., Verma, D.K., Arain, M.A., Finkelstein, M.M.: Correlation of nitrogen dioxide with other traffic pollutants near a major expressway. Atmospheric Environment **42**(2), 275–290 (2008). https://doi.org/10.1016/j.atmosenv.2007.09.042
4. EPA: Nitrogen dioxide (no2) air pollution (2024), `https://www.epa.gov/no2-pollution`, accessed 4-7-2024
5. ESA: Proba-v super resolution. `https://earth.esa.int/eogateway/missions/rapideye` (2021)
6. ESA: Rapideye. `https://earth.esa.int/eogateway/missions/rapideye` (2024)
7. ESA: S2 mission. `https://sentiwiki.copernicus.eu/web/s2-mission` (2024)
8. ESA: Sentinel-2. `https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2` (2024)
9. ESA: Sentinel-5p. `https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P` (2024)
10. ESA: Worldview-2. `https://earth.esa.int/eogateway/missions/worldview-2` (2024)
11. Galar, M., Sesma, R., Ayala, C., Aranda, C.: Super-resolution for sentinel-2 images. ISPRS Journal of Photgrammetry and Remote Sensing **146**, 95–102 (2019). https://doi.org/10.5194/isprs-archives-XLII-2-W16-95-2019
12. Gouvine, G.: Ninasr: Efficient small and large convnets for super-resolution. `https://github.com/Coloquinte/torchSR/blob/main/doc/NinaSR.md` (2021). https://doi.org/10.5281/zenodo.4868308
13. Gouvine, G.: Super-resolution networks for pytorch. `https://github.com/Coloquinte/torchSR` (2021). https://doi.org/10.5281/zenodo.4868308

14. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016). https://doi.org/10.1109/CVPR.2016.90
15. Huang, Y.K., Luvsan, M.E., Gombojav, E., Ochir, C., Bulgan, J., Chan, C.C.: Land use patterns and so2 and no2 pollution in ulaanbaatar, mongolia. Environmental Research **124**, 1–6 (2013). https://doi.org/10.1016/j.envres.2013.02.006
16. Khajeamiri, Y., Sharifi, S., Moradpour, N., Khajeamiri, A.: A review on the effect of air pollution and exposure to pm, no2 , o3 , so2 , co and heavy metals on viral respiratory infections. Journal of Air Pollution and Health **5**(4) (2021). https://doi.org/10.18502/japh.v5i4.6445
17. Lamsal, L.N., Martin, R.V., Parrish, D.D., Krotkov, N.A.: Scaling relationship for no2 pollution and urban population size: A satellite perspective. Environmental Science Technology **47**(14), 7855–7861 (2013). https://doi.org/10.1021/es400744g
18. Lanarasa, C., Bioucas-Dias, J., Galliania, S., Baltsaviasa, E., Schindler, K.: Super-resolution of sentinel-2 images: Learning a globally applicable deep neural network. ISPRS Journal of Photgrammetry and Remote Sensing **146**, 305–319 (2018). https://doi.org/10.1016/j.isprsjprs.2018.09.018
19. Lancet: Gbd 2017 risk factor collaborators (2018). global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the global burden of disease study 2017 **392**(10159) (2017). https://doi.org/10.1016/S0140-6736(18)32225-6
20. Liu, N., Ma, R., Wang, Y., Zhang, L.: Inferring fine-grained air pollution map via a spatiotemporal super-resolution scheme. UbiComp/ISWC pp. 498–504 (2019). https://doi.org/10.1145/3341162.334560
21. Razzak, M.T., Mateo-García, G., Lecuyer, G., Gómez-Chova, L., Gal, Y., Kalaitzis, F.: Multi-spectral multi-image super-resolution of sentinel-2 with radiometric consistency losses and its effect on building delineation. ISPRS Journal of Photogrammertry and Remote Sensing **195**, 1–13 (2023). https://doi.org/10.1016/j.isprsjprs.2022.10.019
22. Rowley, A., Karakus, O.: Predicting air quality via multimodal ai and satellite imagery. Remote Sensing of Environment **293** (2023). https://doi.org/10.1016/j.rse.2023.113609
23. Scheibenreif, L., Mommer, M., Borth, D.: Toward global estimation of ground-level no2 pollution with deep learning and remote sensing. IEEE Transactions on Geoscience and Remote Sensing **60**, 1–14 (2022). https://doi.org/10.1109/TGRS.2022.3160827
24. Scheibenreif, L., Mommert, M., Borth, D.: A novel dataset and benchmark for surface no2 prediction from remote sensing data including covid lockdown measures. IEEE International Geoscience and Remote Sensing Symposium IGARSS pp. 8364–8367 (2021). https://doi.org/10.1109/IGARSS47720.2021.9554037
25. Sdraka, M., Papoutsis, I., Psomas, B., Vlachos, K., Ioannidis, K., Karantzalos, K., Gialampoukidis, I., Vrochidis, S.: Deep learning for downscaling remote sensing images: Fusion and super-resolution. IEEE Geosciences and Remote Sensing Magazine **10**(3), 202–255 (2022). https://doi.org/10.1109/MGRS.2022.3171836
26. Shermeyer, J., Etten, A.V.: The effects of super-resolution on object detection performance in satellite imagery. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops pp. 1432–1441 (2023). https://doi.org/10.1109/CVPRW.2019.00184

27. Sorek-Hamer, M., Pohle, M.V., Sahasrabhojanee, A., Asanjan, A.A., Deardorff, E., Suel, E., andKamalika Das, V.L., Oza, N.C., Ezzati, M., Brauer, M.: A deep learning approach for meter-scale air quality estimation in urban environments using very high-spatial-resolution satellite imagery. Atmosphere **13**(5), 696 (2022). https://doi.org/10.3390/atmos13050696
28. Valsesia, D., Magli, E.: Permutation invariance and uncertainty in multitemporal image super-resolution. IEEE Transactions on Geosciences and Remote Sensing **60**, 1–12 (2022). https://doi.org/10.1109/TGRS.2021.3130673
29. Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., Fu, Y.: Image super-resolution using very deep residual channel attention networks. ECCV 2018 (2018). https://doi.org/10.48550/arXiv.1807.02758
30. Zhu, S., Xu, J., Fan, M., Yu, C., Letu, H.: Estimating near-surface concentrations of major air pollutants from space: A universal estimation framework lapso. IEEE Transactions on Geoscience and Remote Sensing **61**, 1–11 (2023). https://doi.org/10.1109/TGRS.2023.3248180