RESEARCH ARTICLE

Downscaling Surface Albedo to Higher Spatial Resolutions With an Image Super-Resolution Approach and PROBA-V Satellite Images

SAGTHITHARAN KARALASINGHAM[1,2], (Graduate Student Member, IEEE),
RAVINESH C. DEO[1,2], (Senior Member, IEEE), **DAVID CASILLAS-PÉREZ**[3],
NAWIN RAJ[1], **AND SANCHO SALCEDO-SANZ**[1,4]

[1]School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield Central, QLD 4300, Australia
[2]Centre for Applied Climate Science, University of Southern Queensland, Toowoomba, QLD 4350, Australia
[3]Department of Signal Processing and Communications, Universidad Rey Juan Carlos, Fuenlabrada, 28942 Madrid, Spain
[4]Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, 28805 Madrid, Spain

Corresponding author: Ravinesh C. Deo (ravinesh.deo@usq.edu.au)

ABSTRACT For bifacial solar photovoltaic panels, surface albedo plays a crucial role in estimating the radiant energy. Since land surfaces are heterogeneous, the actual albedo of the surface where the solar photovoltaic panel is placed can vary widely and its temporality and sparsity present a significant challenge for renewable energy engineers. This paper develops a new image super-resolution deep learning model based on convolutional neural network to generate high resolution spatial representations of surface albedo from coarse resolution remote sensing-based data. For selected Australian locations, we generated a higher resolution surface albedo using imagery from PROBA-V/SPOT Earth Observation satellites. We proposed a Deep Downscaling Spectral Model with Attention (DDSA) with the capability of processing 10-day albedo images captured at a relatively low ($\approx$ 1 km) resolution. The proposed DDSA was then applied to downscale observed surface albedo and generate predicted albedo at 500 m, 333 m and 250 m resolutions. The proposed model was benchmarked with alternative deep learning, super-resolution approaches: Super-Resolution Convolution Neural Network (SRCNN), Enhanced Deep Super-Resolution network (EDSR), Efficient Sub-Pixel Convolutional Neural Network (ESPCN) and Residual Dense Network (RDN). The results showed that the proposed DDSA model outperformed all comparative models in terms of the mean square error (MSE) $\approx$ 0.0041, signal-to-noise ratio ($PSNR$) $\approx$ 39.471, Structural Similarity Index ($SSIM$) $\approx$ 0.999 *vs.* an MSE $\approx$ [0.0140-0.0387], PSNR $\approx$ [29.761-33.850], SSIM $\approx$ [0.9994-0.999]). We also cross-validated the downscaled images with satellite imagery and ground-based observations, which reaffirmed the proposed DDSA model's ability to produce high resolution surface albedo maps and its potential applications for granular scale tracking and mapping solar energy where bifacial solar photovoltaic panels are placed.

INDEX TERMS Surface albedo downscaling, image super resolution, depth-wise separable convolution, bifacial solar photovoltaic system.

I. INTRODUCTION

As global efforts to achieving the formidable goal of net zero emission by 2050 accelerates, sustainable production and consumption of energy has come under intense focus. Responsible for more than 75% of greenhouse gas emissions globally, a quadrupling of solar and wind capacity annually on 2020 figures is required to achieve sustainable energy production by 2050 [1]. Similarly, our cities where more than 66% of the global population live [2] and currently

The associate editor coordinating the review of this manuscript and approving it for publication was Okyay Kaynak.

consume more than 78% of energy [3], curbing the effects of land usage change on local climate, and hence on energy consumption [4], has become a matter of urgency. Influencing both the production and consumption of energy, the role of surface albedo is now gaining attention [5], [6]. As an indicator of the radiative surface energy and its responsiveness to the land surface characteristics, surface albedo varies both in space and in time. In its nascent stage, the harvesting of this solar radiative surface energy through innovations in solar technology, such as bifacial solar panels [7], has produced increases in solar energy yields of up to 40% when compared with the equivalent monofacial panels from simulation studies [8], [9]. At the same time, curbing the effects of radiative surface energy through greening of urban environments and cooling measures is set to reduce building cooling energy requirements, which currently amounts to 22% of electricity consumption [10], [11]. Despite the dual role of surface albedo, both as a novel energy source and also as a driver of energy consumption, the inability to represent its variability at the degree of temporal and spatial precision required for critical applications such as energy gains and energy consumption modelling is proving to be a major hindrance for its adoption [12]. Therefore, developing novel means of modelling surface albedo, especially at a granular spatial scales, to harness remote sensing capabilities is valuable to increase the solar energy intensity per square meter, thereby enabling cities to mitigate its negative impact and fulfill net-zero commitments.

Identified by the Global Climate Observing System (GCOS) as an Essential Climate Variable (ECV) [13], surface albedo represents the solar radiation reflected by the surface of the Earth. It can be broadly classified into black-sky albedo (the focus of this paper), and white-sky albedo. The former is measured when the sun is at apogee with the assumption that all energy is directly received from the sun, while the latter assumes a part of the energy received through diffusion [14]. Influencing the biophysical state of the earth, albedo, is a unit-less quantity that ranges from 0 to 1 with 0 = perfect absorbance, and 1 = perfect reflectance. As a key driver in energy exchange between the surface and atmosphere, mapping of surface albedo at a granular scale from a nominal 4 km (National Solar Radiation Database) spatial resolution has a strong potential to provide precise albedo inputs for bifacial energy gain modeling of solar farms [15], [16], [17]. This would help accelerate the uptake of bifacial photovoltaic power from the present 20% market share [8] and spur innovative energy generation models such as AgriPV [8]. Higher resolution albedo data can also help urban planners and policy makers to device precise interventions to mitigate the undesirable effects of surface albedo on energy consumption, as the anthropogenic land use effects continues to accelerate. The objective of this research is to develop a novel deep downscaling approach utilising image-super resolution method with an aim to generate higher resolution black-sky albedo observations for mainland Australia. The outcomes are expected to benefit energy modelling applications dependent on high spatial precision surface climatic observations.

II. BACKGROUND AND RELATED WORK

Several studies have attempted to generate granular scale atmospheric features, at a wide range of spatial scales such as from 10 km low resolution maps to 1.25 km high resolution maps. The study in [18] evaluated Ordinary Least Squares, Elastic-Net, State Vector Machines (SVM) and Bias Corrected Spatial Disaggregation (BCSD) against Multitask Sparse Structure Learning (MSSL), BCSD coupled with MSSL and CNN to downscale daily precipitation datasets. Their study found that a linear methods outperformed state-of-the-art machine learning models in capturing anomalies including extreme events or large-scale climate shifts. In a study on uncertainties inherent in statistical downscaling models [19], the authors found that an Artificial Neural Network (ANN) model performed poorly compared with Weather Generator and Statistical Downscaling (SDSM) method evidenced by a greater variability and uncertainty in estimated means and variances of temperature and precipitation. The works [20] and [21] downscaled global climate model predictors to conclude that their Bayesian method produced a high accuracy in downscaling large-scale variables. Adopting a stacked super-resolution image processing method for statistical downscaling of Earth System Model (ESM) in NASA's Earth Exchange research, the study of [18] found that deep-SD model outperformed ANN, SVM, BCSD and SRCNN model. The study in [22] utilized the U-Net CNN algorithm for hyper-local precipitation forecasting, showing that their model outperformed the High-Resolution Rapid Refresh (HRRR), which was a numerical climate model for a 1-hr window forecast. Although these studies have provided a strong promise of deep learning in hyper localisation of precipitation against parametric models, the utility of image-super resolution approach against ground observations of ECV (e.g., black-sky albedo) is yet to be fully explored. An augmentation of the predictive skill of such models with applications in denoising approach and attention mechanisms, as utilised in this paper to achieve a good level of spatial precision competitive to ground-based observations, is a useful way to support critical applications in the solar energy industry and global warming mitigation efforts.

Despite being able to simulate macro-scale climatic phenomenon, the current climate models are somewhat unable to represent the ground conditions at high spatial, temporal and probabilistic precision as required in critical solar energy or other applications. The lack of granularity in spatial datasets either from ground observations or remote sensing satellites and associated uncertainties when attempting to downscale such variables is particularly challenging [23]. This is especially true when granular scale predictions are heavily relied upon to adapt the critical infrastructure, such as weather dependent solar power generation or urban scale transformations, to address the energy demand variations caused by

biophysical drivers such as surface albedo. Added to such uncertainties is the sparsity (or lack) of land-based radiation monitoring capabilities, or the inaccessibility of ground monitoring infrastructure, which also means that ground-based observations generally lack the required granularity to be widely adopted.

The above-mentioned challenges are currently being addressed with better remote sensing capabilities such as utilising satellite-based data platforms. Several studies such as [24], [25], [25], [26], [27] coupled remote sensing variables with machine learning methods to reliably estimate the global solar radiation. In fact, the study of [24] and [25] has explored both solar and wind energy resources with Communication, Ocean and Meteorological Satellite (COMS) Meteorological Imager (MI) geostationary satellite and numerical weather prediction reanalysis datasets where a pixel-based physical model was optimized and a cloud masking was performed to discriminate between the clear and the cloudy area. The cloud simulations were also performed using a cloud factor and clear areas were studied by the atmospheric parameterization method to estimate the surface solar radiation compared with pyranometer observations.

Despite advances in remote sensing platforms, the temporal and spatial coverage and regional data availability imposes restrictions on surface albedo modelling. Though Landsat-8 has a higher spatial accuracy, its 16-day revisit times limit its temporal representation. Similarly, Sentinel-2A and Sentinel-2B launched in 2016 and 2017 provides higher spatial resolution from 5 to 10 days revisit times. Yet, the acquisition capacity was limited over South-eastern Australia in the early phases of operationalisation, increasing revisit times to between 10-20 days, limiting the length of the data available for certain regions [28]. In recent years, new research efforts by NASA have focused on harnessing recent advances in sensor capability from disparate remote sensing platforms, Landsat and Sentinel-2 to generate seamless timeseries of satellite observations. The ongoing Harmonized Landsat and Sentinel-2 (HLS) project develops a series of algorithms for cross-sensor atmospheric correction, cloud masking, view geometry and spectral adjustments to derive surface reflectance of higher temporal and spatial resolution for recent years [29]. With variations within 11% of the MODIS 5km surface reflectance product is hampered by sub-optimal cloud masking.

However, both energy modelling and climate mitigation efforts stands to benefit from the development of super-resolution-based approaches for the harmonisation of spatial resolution of surface albedo data from existing coarse sensor platforms. It would provide the ability to model surface albedo over longer time series which would enhance trend-driven decision-making. Further, it would enable the use of multiple remote sensing platform to pick highly representative surface albedo estimations or used in combination in ensemble modelling for more accurate estimations. In addressing this gap, we consider, SPOT/PROBA-V with

similar orbital characteristics to Sentinel-2, which provides a 10-day composite at 1 km resolution since 1998 [30]. Launched by the European Space Agency (ESA) the onboard sensors were specifically tasked with the mapping of land use and vegetation growth every 2 days. Capable of sensing across spectral bands blue, red, Near-Infrared (NIR) and Shortwave Infrared (SWIR) the sensors provide a ground sampling distance of 200 m at nadir over SWIR and 100 m at nadir over visual and NIR bands. It has demonstrated comparative performance to MODIS and validated against the Surface Radiation Budget (SURFRAD) Network, European Eddy Fluxes Database Cluster (EFDC) and FluxNet [30].

Since each pixel of this remotely sensed multi-spectral imagery is expected to reflect the spectral characteristics that capture the interactions between climate and physical surface characteristics, in this paper, we have advanced the previous research adopting deep learning methods to generate surface albedo maps at granular spatial scales. The research performed in our paper aims to test the proposed downscaling method across the vast surface area such as the Australian mainland, which can be used as climate-based inputs into solar energy modelling or mapping exercise, especially where higher spatial resolution surface albedo data are required.

To build upon the previous research in computer vision [31], as well as super resolution of natural images with Deep Convolutional Neural Network (DCNN) methods [32] and deep learning statistical downscaling methods [18], we couple the remote sensing data with deep learning to capture relationships between spectral observations at high spatial scales in this study. Proposed by [31], image super-resolution approach employ the DCNNs to learn the mapping between low and high-resolution sky images. It operates by utilising a given low-resolution image to reconstruct features synonymous with a high-resolution image that is closer to ground truth. The steps of learning the mapping involves image patch extraction, learning non-linear mappings and reconstruction of learnt patches into a single high-resolution output. As the study of [31] states, which is also evidenced by many other super-resolution methods later e.g., [33], modifications in the model architecture such as up-sampling, improved network design, enhanced learning approaches and feature enhancement methods are warranted to elevate the predictive skill of a basic super resolution framework that can be tailored to address domain-specific applications [33].

This study, therefore, focuses on developing image super resolution approaches to enhance real world imagery and attain higher perceptual accuracy for the surface albedo mapping. To the best of the authors' knowledge, this paper is the first of its kind to introduce a novel deep downscaling approach (referred to as DDSA hereafter) to this end. As shown in Figure 1, the proposed DDSA model extends the deep image-super resolution model by using the Depth wise Separable Convolution, Residual in Residual dense blocks (RRDB) and Convolutional Block Attention (CBAM)
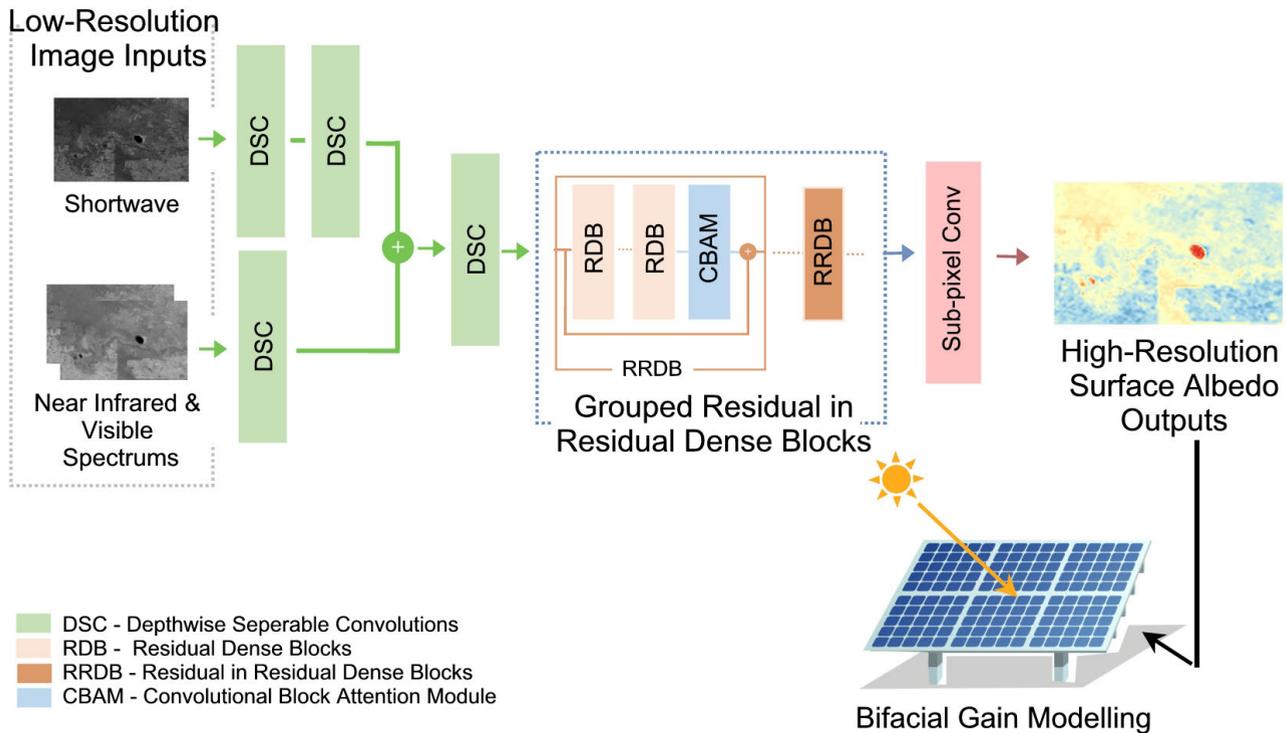
FIGURE 1. Architecture of proposed Deep Downscaling Spectral model with Attention (DDSA) that features Depth Separable Convolution (DSC) Block, Residual in Residual Dense Block (RRDB) and Convolutional Block Attention Module (CBAM) used to develop high spatial resolution surface albedo maps for solar energy applications.

to enhance its effectiveness in efficiently downscaling remote sensed black-sky albedo imagery. Our primary goal is to perform the downscaling of surface albedo from the coarse spatial resolution (e.g. 1 km) images to a more granular scale (e.g. 500 m, 333 m, and 250 m) resolutions. Driven by the intuition that spatially and temporally consistent secondary images could be utilized to learn the missing feature representations arising from casual factors, the proposed DDSA is developed to learn from multiple images representing different spectral bands to extract the features maps that are then used to enhance the downscaled imagery.

More precisely, the study harnesses the decomposed spectral imagery from the Near-Infrared ($[0.7\mu m\text{-}4\mu m]$), Visible ($[0.4\mu m\text{-}0.7\mu m]$) and total shortwave ($[0.3\mu m4\mu m]$) bands as a proxy to represent biophysical parameters of surface classes which gives rise to the variations in radiative surface energy. Further, we explicitly avoid the use of auxiliary datasets on land cover and instead rely on deep learning implicit representation of the land cover from the raw spectral imagery. In this regard, the integration of grouped RRDB blocks is performed to facilitate the learning of hierarchical features while mitigating the noise, contributing to an enhanced quality output. The addition of CBAM further refines the learning, to focus on essential feature representations, thereby enhancing the high-resolution representation. The computational intensity of the end-to-end super resolution process is also addressed through the replacement of standard convolutions with computationally less costly i.e.,

the Depth wise Separable Convolutions and the learning of mapping between low resolution and high-resolution imagery in the low-resolution space with Sub-pixel Convolutions.

Therefore, the main contributions of this work are: (i) the development of deep downscaling approach based on image-super resolution to harness remote sensed coarse spectral imagery, learn across spectra features and generate the higher resolution black sky albedo; (ii) the combination of de-noising and attention mechanisms to mitigate and reduce the noise contributions that usually occur at higher resolutions while enhancing the predictive accuracy of the final albedo map; (iii) the correct utilization of computationally efficient approach for end-to-end super resolution process, while outperforming the competing benchmark models: Super-Resolution Convolution Neural Network (SRCNN), Enhanced Deep Super-Resolution network (EDSR), Efficient Sub-Pixel Convolutional Neural Network (ESPCN) and Residual Dense Network (RDN); and (iv) the elicitation of scientific utility of newly proposed deep image super resolution approach for deep downscaling remote sensed sky imagery to enhance the spatial granularity.

The remainder of the paper has been structured in the following way: next section presents the proposed DDSSA model and its building blocks. Section IV describes the materials and methods used to assess the performance of the proposed approach, including the study area, model design and performance criteria metrics. Section V describes the experimental part of the work, where we evaluate the

proposed model in real data in Australia. Section VI closes the paper with some final conclusions, remarks and future lines of research.

III. THEORETICAL OVERVIEW

We now present the details of the proposed image-super resolution DDSA model shown in Figure 1. The proposed model is based on four distinct methods: Depth wise Separable Convolution, Residual in Residual dense blocks (RRDB), Convolutional Block Attention Module (CBAM) and sub-pixel convolution algorithms. For theoretical details of all benchmark models, readers can consult references elsewhere: e.g., SRCNN [31], EDSR [32], ESPCN [34] and RDN [35].

A. DEPTH WISE SEPARABLE CONVOLUTION

It is imperative to mention that Convolutional Neural Networks (CNN) as the standard convolutions made up of convolution layers that employ kernels to extract features to build a compact representation in form of feature maps [36]. In such convolution algorithms, the underlying channel (or the depth) and the spatial computation is carried out simultaneously to produce an output feature map as shown in Equation 1 resulting in a rather computationally expensive [37] [38] or a time intensive process [39]. Assuming that an image is a multidimensional matrix (I) with a channels or depth (C) of size $W \times H \times C$, applying a filter (K) of size $M \times M \times C \times N$ on an input feature (F) of size $D \times D \times C$, we can produce a feature map (O) of size $D \times D \times N$ [38], [40]:

$$O_{k,l,n} = \sum_{i,j,m} K_{i,k,m,n} \cdot F_{k+i-1, \, l+j-1, \, m} \quad (1)$$

By contrast, in a depth wise separable convolution algorithm, the process is usually decomposed into a sequential step of depth wise convolution which is performed on each channel followed by a point wise convolution which combines the channel wise feature maps into a combined output [41]. Consistent with our aim in this paper, this can now occur at a lesser computational cost due to the reduced number of parameters [37] [42]. Accordingly, firstly by applying filter (K) of size $M \times M$ depth wise on input feature (F) and channel wise (C), we obtain

$$\widehat{O}_{k,l,n} = \sum_{i,j} \widehat{K}_{i,k,m,n} \cdot F_{k+i-1, \, l+j-1, \, m} \quad (2)$$

By applying a filter (K) of size 1×1, which is performed point-wise on the depth wise output from the above, one obtains

$$O_{k,l,n} = \sum_{m} \widetilde{K}_{m,n} \cdot \widehat{O}_{k-1,l-1,m} \quad (3)$$

It is noteworthy that the computational complexity can actually make it impractical to adopt any image-super resolution method in resource-constrained environment [43]. Therefore, the significantly reduced computational cost and associated time efficiency attained in the proposed DDSA approach using depth wise separable convolutions is of great

significance in deep learning image-super resolution application for future solar energy monitoring, forecasting and feasibility studies. However, the use of the depth wise separable convolution is also of significance as hypothesized by [41] where combining depth and spatial correlations is expected to be far less efficient than a similar process carried out independently. The underlying premise is that the independent process of depth correlation and spatial correlation is expected to lend itself to the learning of high level and low-level features separately and the sharing of low level features across different image domains in visual space [40]. We therefore leveraged this in spectral space to feed through spectral wise broadband albedo imagery that encompasses different surface characteristics across vast geography of Australian mainland. Hence, our method helps to learn the high and the low-level features as well as combining the cross spectral features for improved downscaling of surface albedo.

B. RESIDUAL IN RESIDUAL DENSE BLOCKS

The proposed DDSA model is configured in such a way that it uses residual dense blocks as a means to enhance the accuracy of the downscaled high resolution surface albedo map. As deeper networks are employed for image restoration enhancement and denoising tasks, the lack of hierarchical features from low resolution inputs can be a major shortcoming that contributes to poor restoration capability of models [35] [44]. We therefore address this shortcoming by combining residual and dense layers into residual dense blocks (RDB) that extracts feature representation locally while learning from preceding RDB layers [35]. The residual in Residual blocks (RRDB) further extends the capability to use RDB blocks of varying depths and adjusting the magnitude of residual contributions through a scaling factor bounded by [0, 1] [45], thereby achieving a higher perceptual quality such as that attained by enhanced super-resolution Generative Adversarial Networks. Similarly, grouped RDB blocks can also achieve higher denoising of the images [44] so the proposed DDSA model, we have used such image restoration capabilities to mitigate the noise and further improve the feature representations locally through grouped RRDB blocks with attention methods.

C. CONVOLUTIONAL ATTENTION BLOCKS

The design of the DDSA model is further inspired by the role of attention in prioritizing visual processing of essential information in visual space [46] given that attention mechanism can enable feature representations through refinements, and discarding the less important information [47] [48] [49]. The Convolutional Block Attention Module (CBAM) used in designing the proposed DDSA model has enabled feature refinements by extracting essential information on spatial and channel dimensions [48]. We therefore exploited the added advantage of lightweight nature of the CBAM especially in terms of the computational costs and by adding the CBAM

after the RDB blocks in the RRDB layer to refine the spectral features prior to generating the RRDB-based outputs.

D. SUB-PIXEL CONVOLUTION

In traditional deep learning image-super resolution, the mapping between low resolution (LR) input and high resolution (HR) outputs is learnt in the HR space with the help of an upscaled LR image [31]. However, to alleviate the computational complexity, the authors in [34] have proposed a sub-pixel convolution or a specialized case of deconvolution where mapping is learnt in the LR space with the upscaling [HR] carried out in the last layer. The proposed DDSA model therefore adopts a variation of the sub-pixel convolution as proposed in an earlier study [50]. This step is expected to address the checkerboard artifacts while also reducing the computational costs of the super resolution process.

IV. MATERIALS AND METHODS

A. STUDY AREA

The proposed DDSA model was constructed using remotely sensed, 10 day directional-hemispherical reflectance imagery for mainland Australia (25.2744°S, 133.7751°E). The directional-hemispherical reflectance imagery from PROBA-V and SPOT satellites has provides the broadband albedo observations at 1 km resolution across the spectral bands over visible (0.4μm-0.7μm), near-infrared (0.7μm-4μm) and total shortwave (0.3μm-4μm). Launched in 1998 and 2014 respectively, SPOT and PROBA-V provides a global coverage with a synthesized Directional and Hemispherical Albedo product generated at 10 day interval. Spectral band imagery for the summer period from 2001 to 2019 were extracted from biophysical data on earth energy budget provided by Copernicus Global Land Service [51] and incorporated as model inputs.

Ground-based observations (GBOV) were extracted from European Commission Joint Research Centre [52] for Validation. Representing diverse surface characteristics from mixed forests to shrub lands as well as evergreen broadleaf and woody savannas, the study area encompasses 6 micrometeorological monitoring sites of the OzFlux network [53] and global FluxNet which are used as ground-based observation sites in this study. Ground observations from 50-m high tower-mounted albedometer was considered to represent the observations of 500 m radius around the tower according to its Field of View (FOV) [54]. To ensure spatial representativeness of tower based surface albedo observations, as per previous studies into ground validation of MODIS surface albedo (1 km and 500 m products) [55], tower locations with homogenous surface characteristics were chosen except 1 location for comparative purposes. Table 1 shows the geographic surface classification and elevation information for OzFlux micro-meteorological monitoring Tower locations where the proposed DDSA model was validated for surface albedo mapping and Figure 2 plots the validation study sites.
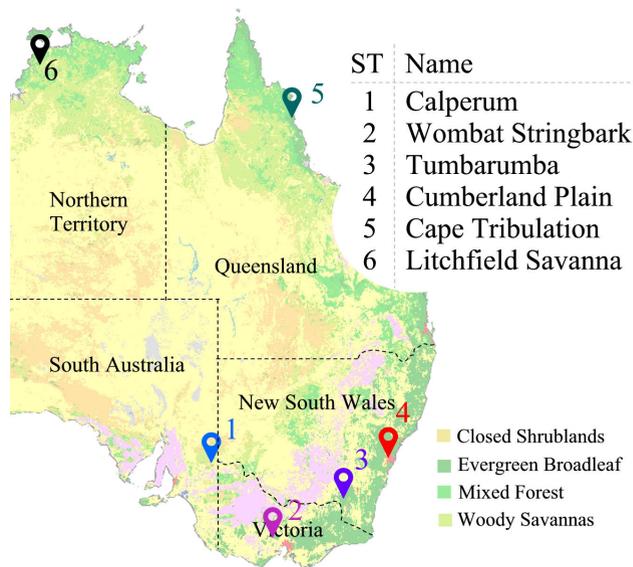

FIGURE 2. A spatial map of OzFlux ground based observation validation (GBOV) towers providing ground truth to evaluate the proposed DDSA model adopted for downscaling surface albedo dataset.

Further validations were carried out with MODIS 500 m surface albedo product (MCD43A3) using bands 1-7 against the PROBA-V/SPOT surface albedo downscaled to 500 m by the DDSA model.

B. MODEL DESIGN

The proposed DDSA, and all benchmark models were developed on Pytorch machine learning framework and trained on dual NVidia RTX 3090 (24Gb) GPU cloud Service. Pytorch provides a low-level Application Programming Interface offering flexibility in experimenting novel algorithmic approaches with a rich set of libraries to tackle application scenarios (https://pytorch.org/). QGIS, a Geospatial data analysis application was used along with Geospatial Data Abstraction Library (GDAL) and netCDF Kitchen Sink (ncks) to process remotely sensed data in netCDF format.

As the primary objective of the proposed DDSA model design was to augment the resolution through a learning of non-linear spectral relationships in surface albedo, we first extracted broadband albedo across spectral bands, visual, near-infrared and shortwaves for Australian mainland and ground-based validation stations. Followed by a conversion to 16-bit tiff images as a lossless image format with GDAL translation, the images were split into training (2001-2010) and validation (2011-2015 and 2017-2019) sets, and used to generate model inferences for 2016 for each ground-based validation locations. To improve the performance of super-resolution method in recovering original HR images, they were subjected to colour space conversion from red-green-blue (RGB) to YCbCr to separate the Luminance (Y) and Chrominance (Cb, Cr) components [56]. Converted images were interpolated using bi-cubic interpolation as a synthetic LR image pair, prior to being normalised in [0, 1]

TABLE 1. Geographical surface classification and elevation information for the OzFlux micro-meteorological monitoring Tower locations where the proposed Deep Downscaling Spectral model with Attention (DDSA) model was validated for surface albedo mapping.

Test Site	Latitude	Longitude	Elevation (m)	Surface Class
Calperum	34° 0′ 7.2″	140° 35′ 13.2″	53	Closed Shrublands
Wombat	37° 25′ 19.2″	144° 5′ 38.4″	112	Evergreen Broadleaf
Stingbark Tumbarumba	35° 39′ 21.6″	148° 9′ 7.2″	1200	Evergreen Broadleaf
Cumberland Plain	33° 36′ 54″	150° 45′ 22.8″	25	Evergreen Broadleaf
Cape Tribulation	16° 6′ 21.6″	145° 22′ 40.8″	570	Mixed Forest
Litchfield Savanna	13° 10′ 48″	130° 47′ 24″	225	Woody Savannas

range as per Equation (4):

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (4)$$

Following data preparation, the proposed DDSA and the benchmark models SRCNN, EDSR and ESPCN together with CNN and RDN with residual dense layers were developed as per model architecture details listed in literature to evaluate the efficacy of the proposed model. Other than the model-specific hyperparameters, a common set of hyperparameters were selected across models, which were as follows:

(a)
1) **Activation Function:** Tangent Hyperbolic (*tanh*) delivered a comparatively better image enhancement through experimentation than Rectified Linear Unit (*ReLU*) or *Leaky Relu* for all except the output layer.
2) **Optimizer:** The Adaptive Moment Estimation (*Adam*) was adopted as an optimizer after comparative testing with Stochastic Gradient Descent (SGD) and AdaBound method.
3) **Hybrid Loss function:** Hybrid loss function [Equation 5] used to jointly minimize Mean Square Error (*MSE*) loss while maximizing Structural Similarity Index (*SSIM*) [57] applied to simultaneously optimise the objectives by minimising MSE and maximising signal-to-noise-ratio (*PSNR*).

Unlike traditional metrics such as *MSE* and closely related *PSNR*, the *SSIM* quantifies the similarity between ground truth and compares an image with emphasis on structural information inherent in respective images [57]. This in fact a perceptual metric quantifying the image quality degradation caused by the processing such as data compression and losses in data transmission, for example, in our case, the downscaling of surface albedo images from 1 km to 250 m resolution. Our experiments noted that the *SSIM* was quite sensitive to distortions in original images while the *MSE* and the *PSNR* presented similar or higher values that did not represent the perceptual image quality [57]. With its scaling parameter α, the hybrid loss function is:

$$\alpha = (1 - \alpha) \cdot MSE + \alpha \cdot SSIM \qquad (5)$$

C. PERFORMANCE CRITERIA

In context of visual modelling frameworks such as image super-resolution used in this study, performance measures utilising *MSE*, *PSNR* and *SSIM* are used to evaluate similarity between predicted and referenced images at high resolutions. We also carried out an ablation study to analyse the contributions of model architecture components to the overall model's performance. Mathematically, the performance evaluation metrics are:

Mean Squared Error, *MSE*:

$$MSE = \sum_{i=1}^{D}(x_i - y_i)^2 \qquad (6)$$

Signal to Noise Ratio, *PSNR*:

$$PSNR = \frac{10 \log_{10}\left(\max\left(\max(x), \max(y)\right)^2\right)}{|x - y|^2} \qquad (7)$$

Structural Similarity Index Measure, *SSIM*:

$$SSIM = \frac{(2\mu_x\mu_y + C_1) + (2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \qquad (8)$$

Note that the mean is represented by μ, C_x is a stability constant when denominator is set to 0, whereas x and y are the observed (lower resolution) and the predicted (higher resolution) image pixels.

V. EXPERIMENTS AND RESULTS

In this section, we now appraise the performance of the newly proposed deep learning image super-resolution approach for downscaling surface albedo at higher spatial resolutions for solar energy and global warming mitigation applications. The proposed DDSA model is quantitatively and qualitatively evaluated to highlight its merits over comparative models (such as SRCNN, EDSR, ESPCN and the RDN method) employed in this problem of deep downscaling of multi-spectral broadband albedo to produce high resolution maps. A comparative assessment on the validation sets are followed by an integrated analysis of visual and quantitative image evaluation metrics on tested data using ground validation stations located at 500 m, 333 m and 250 m resolution.

A comparative evaluation of all prescribed models based on their capability to achieve the lowest *MSE*, highest *PSNR* and *SSIM* shows that the objective model DDSA appears to outperform all of the other models. This yielded an *MSE* ≈ 0.0041, *PSNR* ≈39.471, *SSIM* ≈ 0.999. These contrasted an
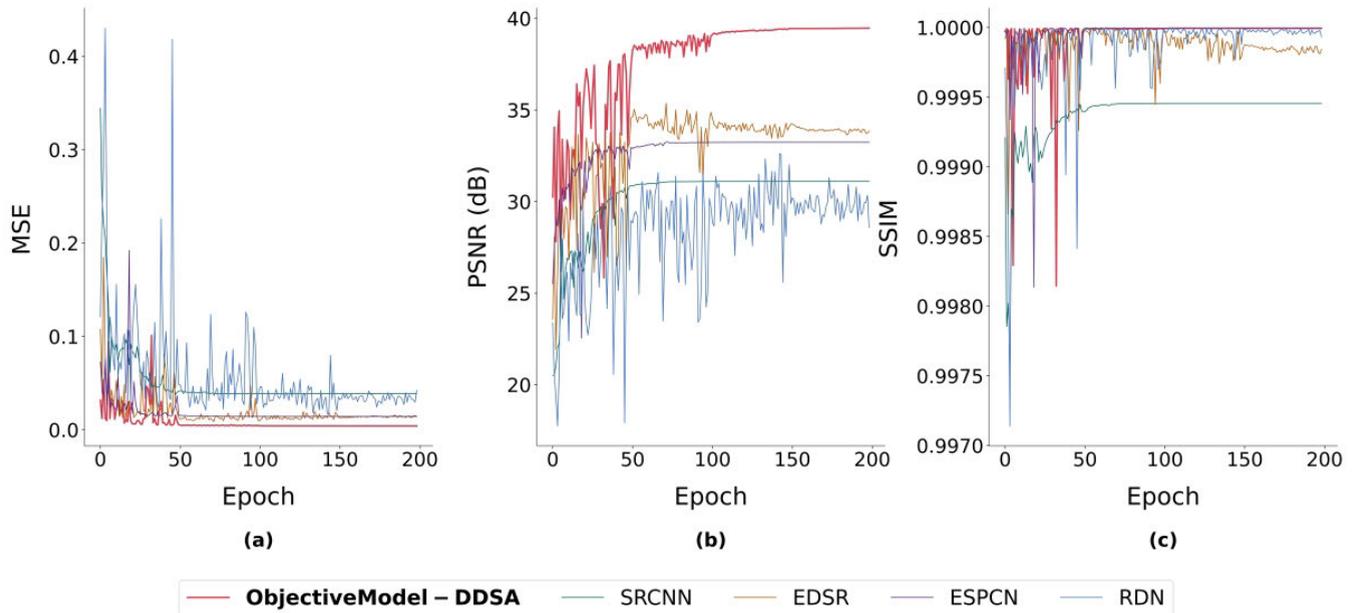
FIGURE 3. Training performance of DDSA model against benchmark models utilizing super-resolution Convolution Neural Networks (SRCNN), Enhanced Deep Super-Resolution Network (EDSR), Efficient Sub-Pixel Convolutional Neural Network (ESPCN) and Residual Dense Network (RDN) used to produce high spatial resolution albedo maps. The mean square error (*MSE*), signal to noise ratio (*PSNR*) and structural similarity index (*SSIM*) are presented for several training epochs.
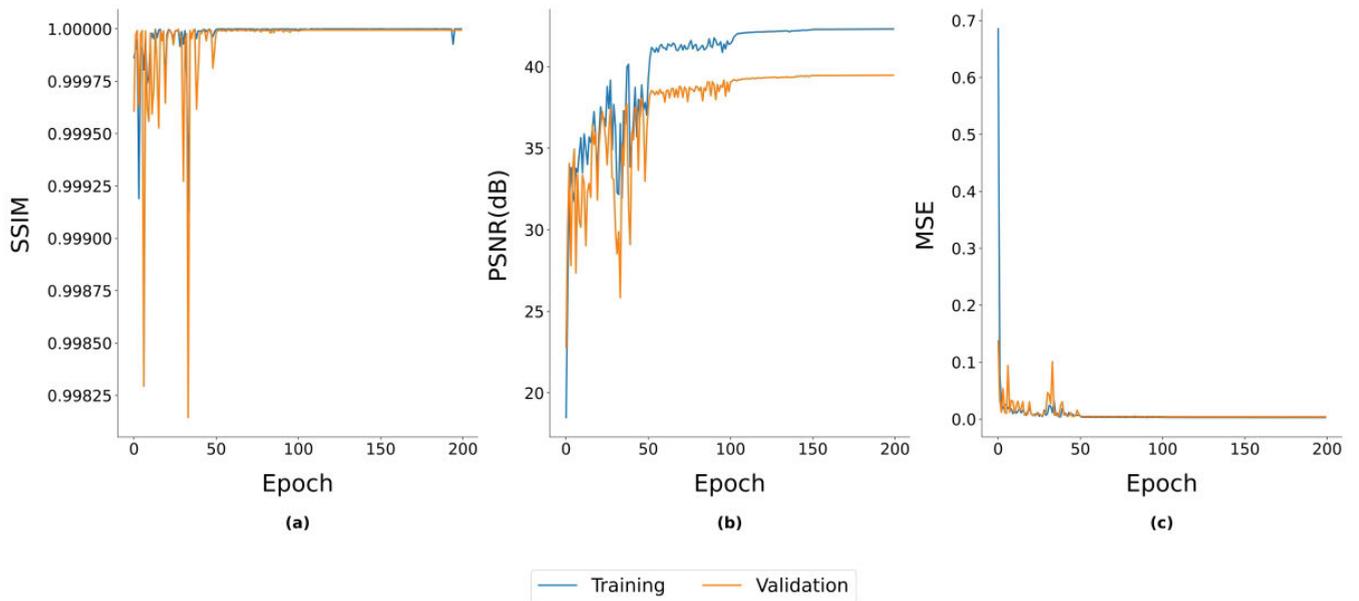

FIGURE 4. Training and validation performance of the proposed DDSA model used to generate high spatial resolution surface albedo maps measured by *MSE*, *PSNR* and *SSIM* vs. epochs.

MSE value $\approx$ [0.0140 - 0.0387], *PSNR* $\approx$ [29.761 – 33.850], and *SSIM* $\approx$ [0.9994 – 0.999]). Figure 3 plots these findings.

In accordance with these findings, the newly proposed DDSA model achieved a sustained high performance after 50 epochs while comparative models require longer training horizons to achieve stable performance as shown in Figure 4.

In translating superior quantitative performance to achieve higher qualitative outcomes, Figure 5 shows the proposed

DDSA model reaching the closest visual representation in respect to ground truth using predictions for OzFlux ground station region, Calperum, Australia as an example. As there is no effective standard for assessment of the perceptual accuracy, we present the comparative visual model predictions for the benchmark models (SRCNN, EDSR, ESPCN and RDN). Notably, the RDN model, despite its lower *MSE* and *PNSR* compared with the objective model, achieved a

(a) Ground Truth   (b) Objective Model (DDSA)   (c) SRCNN

(d) EDSR   (e) ESPCN   (f) RDN
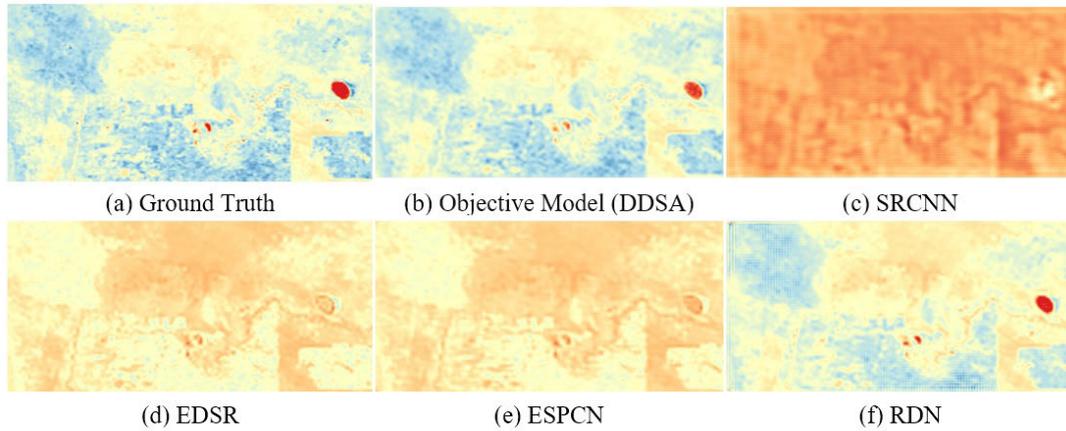
Metrics	Objective Model	Benchmark Models			
	DDSA	SRCNN	EDSR	ESPCN	RDN
PSNR (db)	39.4710	31.1021	33.8504	33.2365	29.7608
SSIM	0.9999	0.9994	0.9998	0.9999	0.9999
MSE	0.0041	0.0387	0.0140	0.01421	0.0341

FIGURE 5. Visual and quantitative model prediction accuracy assessment of the objective DDSA model *vs.* the benchmark models i.e., SRCNN, EDSR, ESPCN and RDN tested for Calperum FluxNet Tower site. *Interpretive note: the optimal method, i.e., the DDSA model, is expected to attain the highest PSNR and/or the SSIM metric and the lowest MSE value.*
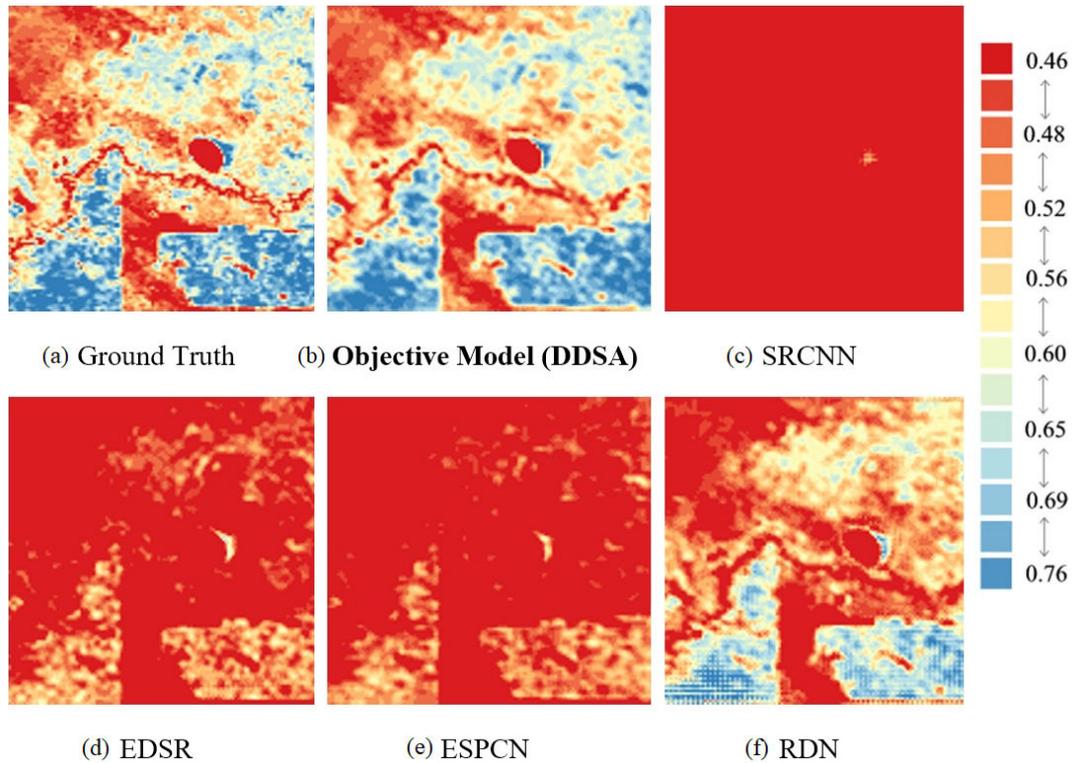

(a) Ground Truth   (b) **Objective Model (DDSA)**   (c) SRCNN

(d) EDSR   (e) ESPCN   (f) RDN

FIGURE 6. Evaluating the proposed DDSA model against the benchmark models across the 16-Quantile values (unit-less) for Calperum FLUXNET Tower site sub-region's remotely sensed Broadband Directional Albedo captured by PROBA-V/SPOT imagery ground truth.

better structural similarity congruent with the ground truth images. The lower PSNR scores despite the higher perceptual performance of RDN is consistent with literature ([58], [59]),

on the effects of brightness and pixel shifts in lowering the score. Further, the high perceptual performance was perhaps due to the architecture of the RDN model enabling it to
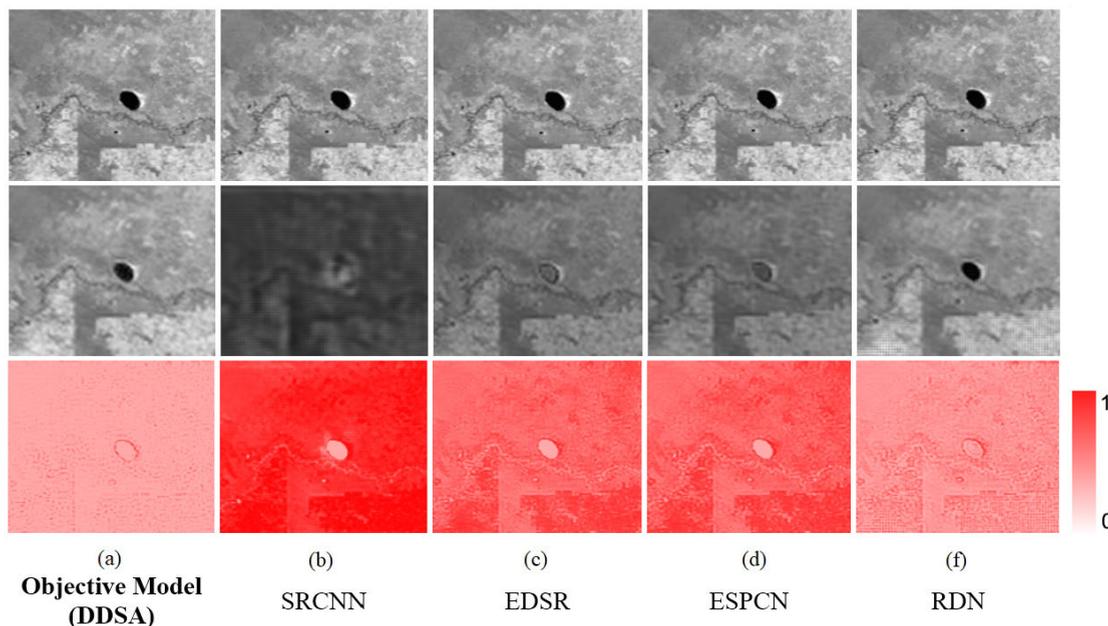
| (a) | (b) | (c) | (d) | (f) |
| Objective Model (DDSA) | SRCNN | EDSR | ESPCN | RDN |

FIGURE 7. Side-by-side comparison of error maps (third row) representing the differences between ground truth (first row) vs model predictions (second row) of the proposed DDSA model, SRCNN, EDSR, ESPCN and RDN models for Calperum FluxNET Tower sub-region site. Lighter shades of red indicate the least difference between ground truth and model predictions. Note that the ground truth remote sensing Broadband Directional Albedo is obtained from PROBA-V/SPOT (first row).
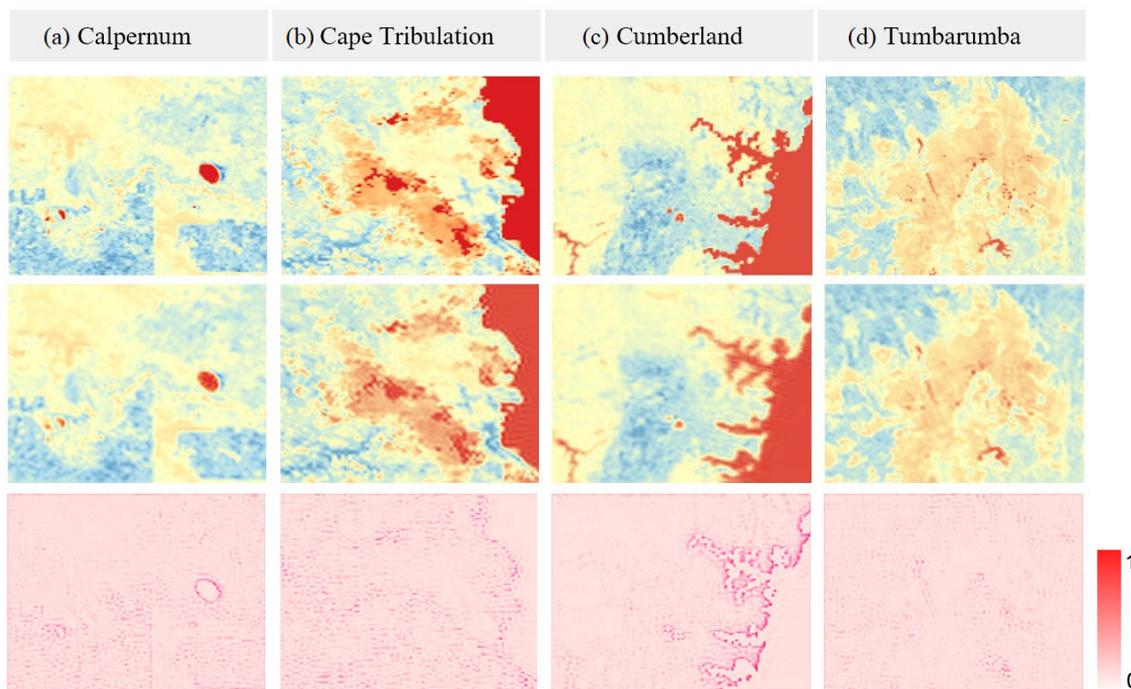

| (a) Calpernum | (b) Cape Tribulation | (c) Cumberland | (d) Tumbarumba |

FIGURE 8. Side-by-side comparison of error maps (third row) representing the differences between ground truth (first row) vs the proposed DDSA model predictions (second row) for FluxNET Tower subregions - a) Calperum, b) Cape Tribulation, c) Cumberland and d) Tumbarumba, Australia. Lighter shades of red indicate the least difference between ground truth and model predictions. Note that the ground truth remote sensing Broadband Directional Albedo is obtained from PROBA-V/SPOT (first row).

harness the residual dense blocks. Hence, it will now be used in further analysis, alongside the objective model developed in this study.

We now revert to the 16-quantile intensity analysis of a portion of ground truth pixel distribution of model predictions across all models to reveal if the best reproduction of ground truth is attained by the DDSA model (Figure 6). It is clear that the objective model closely resembles the ground truth image although the RDN model also appears to generate a close enough image similar to the ground
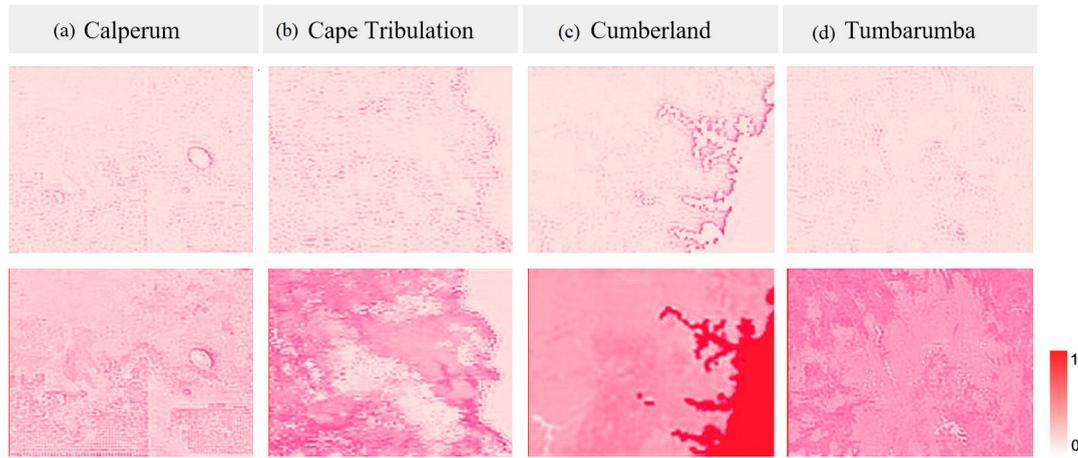
FIGURE 9. Side-by-side comparison of error maps between ground truth and high spatial resolution surface albedo predictions of objective model (DDSA) (first row) and benchmark model RDN (second row) for the FluxNET Tower subregions - a) Calperum, b) Cape Tribulation, c) Cumberland and d) Tumbarumba, Australia. Lighter shades of red indicate the least difference between ground truth and model predictions. Note that the ground truth remote sensing Broadband Directional Albedo is obtained from PROBA-V/SPOT (first row).
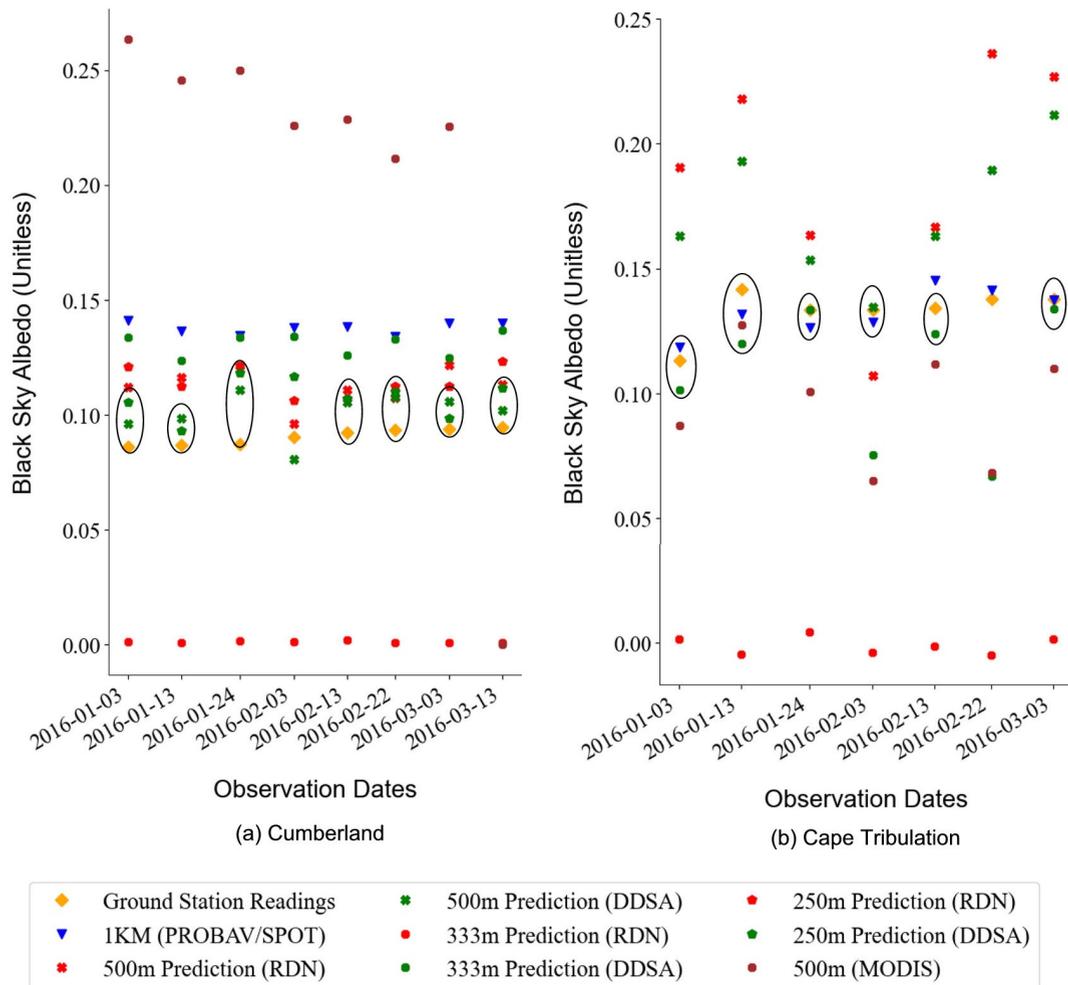

FIGURE 10. Applying the proposed DDSA model predictions of surface albedo contrasted against FluxNet locations: (a) Cumberland and (b) Cape Tribulation, Australia Station predictions and Satellite observations (PROBA-V/SPOT).

truth image. Notably, the quantile plots reveal the transition between pixel subsets in the RDN, which does display grid artefacts in higher value regions while the proposed DDSA model is more effective in removing such artifacts and to further generating a more natural looking predicted image.
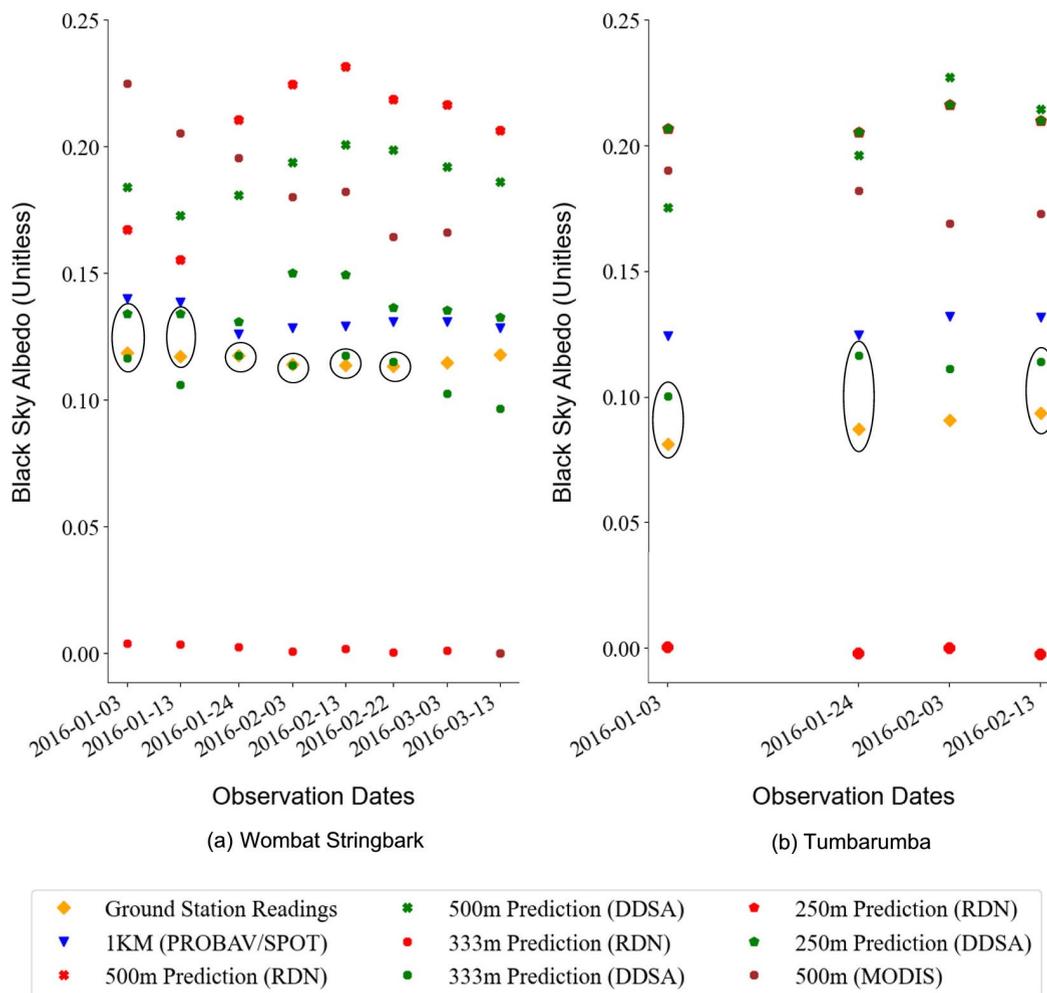
FIGURE 11. Applying the proposed DDSA model predictions of surface albedo contrasted against FluxNet locations: (a) Wombat Stringbark and (b) Tumbarumba, Australia Station predictions and Satellite observations (PROBA-V/SPOT).

We also evaluate all the developed models in terms of errors captured in generating most representative prediction at high spatial resolutions. As a result, we note that the proposed DDSA model produces lowest error in comparison with the other counterpart models. Error maps, as shown in Figure 7 presents the visual distribution of the magnitude of error between the ground truth and the model prediction for OzFlux ground station region of Calperum. A side-by-side comparison reveals that the proposed DDSA model outperforms all, with excellent capability to generate the predictions closer to the ground truth in comparison with the SRCNN, EDSR, ESPCN or the RDN model.

Although producing the next best results, the RDN model is consistent with earlier findings that displayed a larger error possibly due to artefacts as highlighted earlier. A consideration of the proposed DDSA model predictions across different ground-based observation station regions reveals the model's capability to generate consistent and natural-looking map that is closer to ground truth predictions. With the aid of error maps we calculate the predicted error between ground

truth and DDSA predictions at selected stations, Calperum, Cape Tribulation, Cumberland and Tumbarumba encompassing a variety of surface classes (Figure 8). The process was repeated against RDN model predictions. A side-by-side comparison as in Figure 9, shows that in comparison with RDN model, the proposed DDSA model minimises the predictive error quite significantly, across FluxNet Tower sites.

After establishing the fact that the proposed DDSA model outperforms all other models (i.e., SRCNN, EDRS, ESPCN and RDN) in terms of qualitative and quantitative metrics we now assess its predictive capability against the PROBA-V/SPOT observation and ground station readings for summer period in 2016. Contrasting the deep downscaled DDSA model prediction at 500 m, 333 m and 250 m against the ground observations and at 1 km PROBA-V/SPOT for the Cumberland station, both the 500 m and the 250 m predictions respectively are within 0.010 – 0.024 and 0.005 – 0.031 of ground observations [Figure 10a]. In contrast, the PROBA-V/SPOT observations slightly deviated from ground observation results between 0.041 and 0.055.
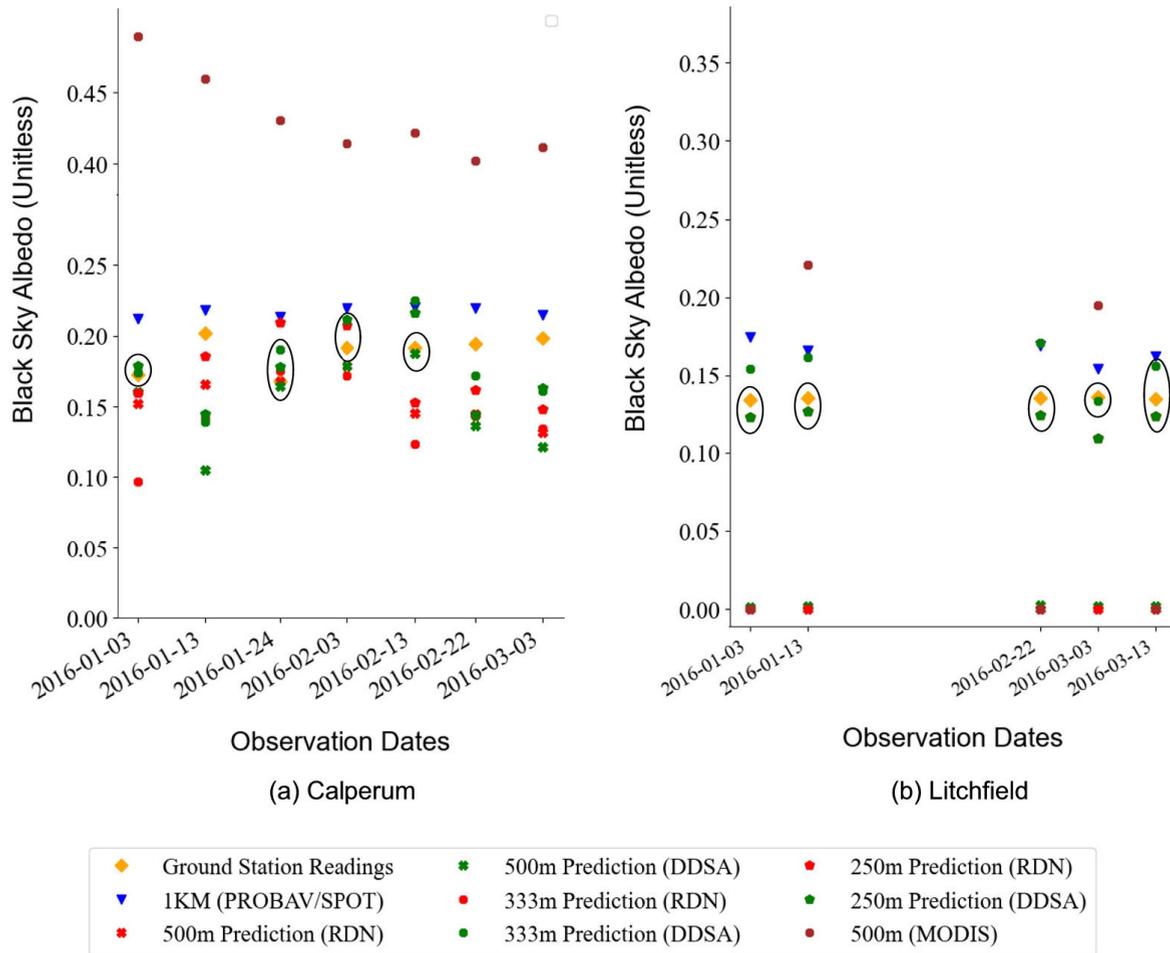
FIGURE 12. Applying the proposed DDSA model prediction of surface albedo contrasted against FluxNet locations (a) Calperum; (b) Litchfield, Australia Station predictions and Satellite observations (PROBA-V/SPOT).

In accordance with Figure 10b, similarly for Cape Tribulation region, the predictions for 250 m resolution achieved the lowest absolute deviation (0.00 – 0.26) while that for the PROBA-V/SPOT deviated between 0.00 – 0.41. Absolute deviations between 0.00 – 0.12 were observed for the case of Wombat Stringbark (Figure 11a) with the 333 m predictions in contrast to PROBA-V/SPOT which deviated between 0.015 – 0.021. For Tumbarumba region [Figure 11(b)] the closest predictions to ground truth are obtained from the predictions at 333 m with absolute deviations between 0.00 – 0.29 while the PROBA-V/SPOT observations deviated further with a range of 0.0 – 0.43.

In contrast, the predictions for Cape Tribulations [Figure 10b], we note a higher absolute deviation for the closest prediction 333 m, while PROBAV/SPOT observations' absolute deviations are 0.00 – 0.011. Similarly, in Calperum [Figure 12a] the model deviations were higher between 0.006 and 0.057 while PROBA-V/SPOT observations were the closest with absolute deviations between 0.016 and 0.04. The 250 m predictions showed the lowest absolute deviations of 0.02 – 0.03 for Licthfield

[Figure 12b] with PROBAV/SPOT observations showing significant divergence.

The highest absolute deviations for Calperum in South Australia, were noted on days the region had experienced temperatures +8 to +9°C higher than the long term average (http://www.bom.gov.au/). A similar pattern was observed for Cape Tribulation, a site with heterogeneous surface characteristics, when the region had experienced higher than usual mean temperatures (http://www.bom.gov.au/).

Further, the DDSA model predictions for surface albedo at 500 m resolution were evaluated against MODIS surface albedo estimates at 500 m resolution. A plot of the absolute deviations (Figure 13) between ground observations and DDSA model predictions vs MODIS surface albedo estimations shows least deviations for DDSA model predictions while MODIS estimates deviated significantly for sites Calperum and Cumberland. Both DDSA model predictions and MODIS albedo estimated showed deviations for both Capte Tribulation with heterogeneous surface characteristics and Wombat Stringbark. A clear trend could not
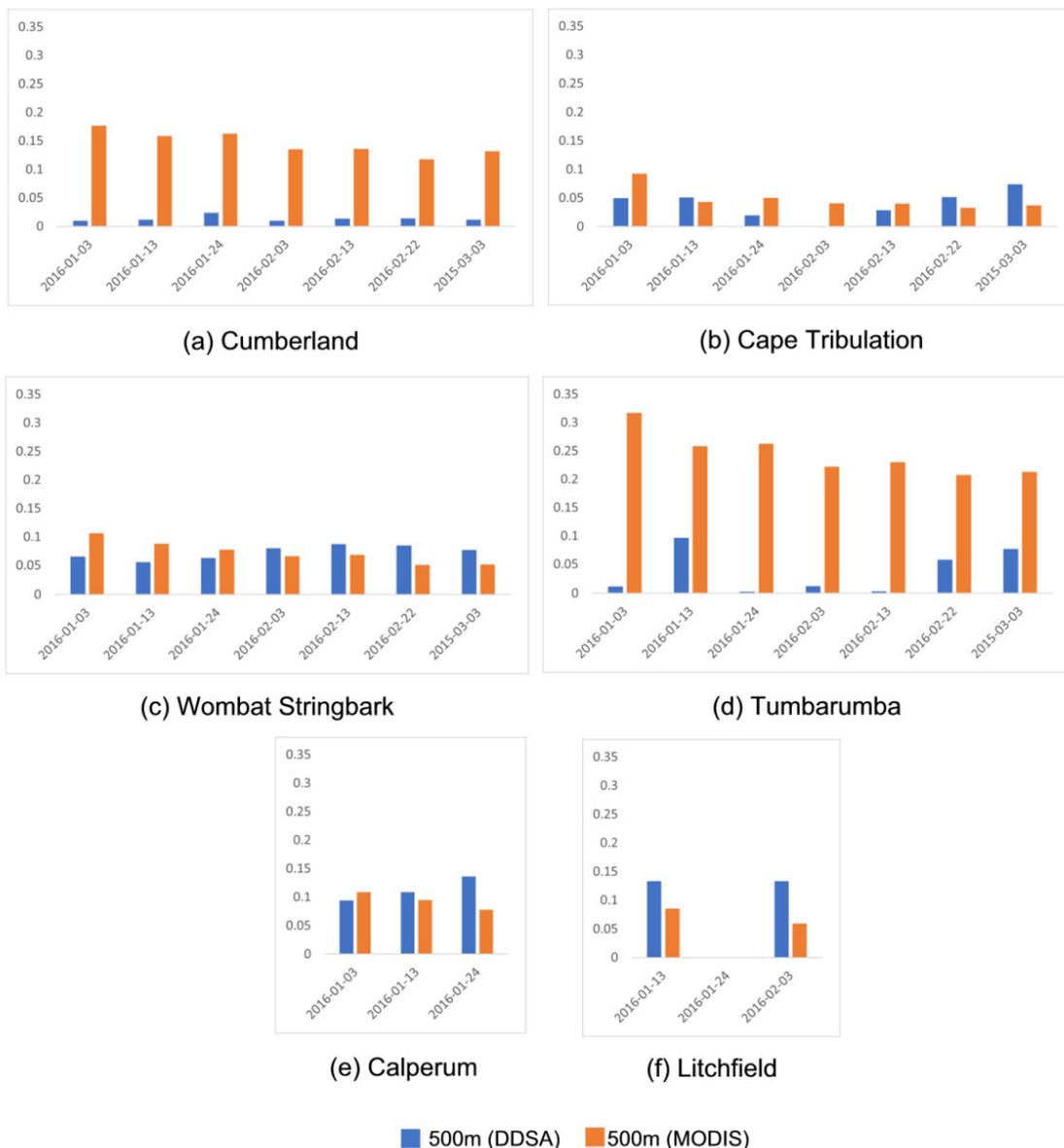
FIGURE 13. Comparison of absolute differences between ground observations and DDSA model downscaled prediction for 500 m and MODIS 500 m surface albedo for FluxNet locations (a) Cumberland; (b) Cape Tribulation; (c) Wombat Stringbark; (d) Tumbarumba; (e) Calperum; (f) Litchfield, Australia.

be established for remaining sites due to the unavailability of cloud free pixel observations for MODIS surface albedo product.

A. ABLATION STUDIES

In this study we assessed the contributions of the different neuronal layers incorporated in DDSA model evaluated through a removal of the convolutional Block Attention Module (CBAM) and the Residual in Residual dense blocks (RRDB) layers with the same data used earlier. The purpose of this part was to investigate the performance of the artificial intelligence system by removing certain components to better understand the contribution of the component to

the overall DDSA system. The performance of the resulting DDSA model was then evaluated using the average value of the *MSE*, *PSNR* and *SSIM* metrics. Figures 14 and 15 shows the ablation study results.

In accordance with ablation studies, we first assessed the proposed DDSA without the CBAM layer to investigate the effect of the attention module. Notably, this change has a major effect on the proposed DDSA model accuracy. It is evident that the removal of attention layer results in the model achieving a lower performance based on *MSE*, *PSNR* and *SSIM* as illustrated in Figure 14 compared with DDSA model trained with CBAM layer. The model with attention achieved a consistent performance at approximately 50 epochs with
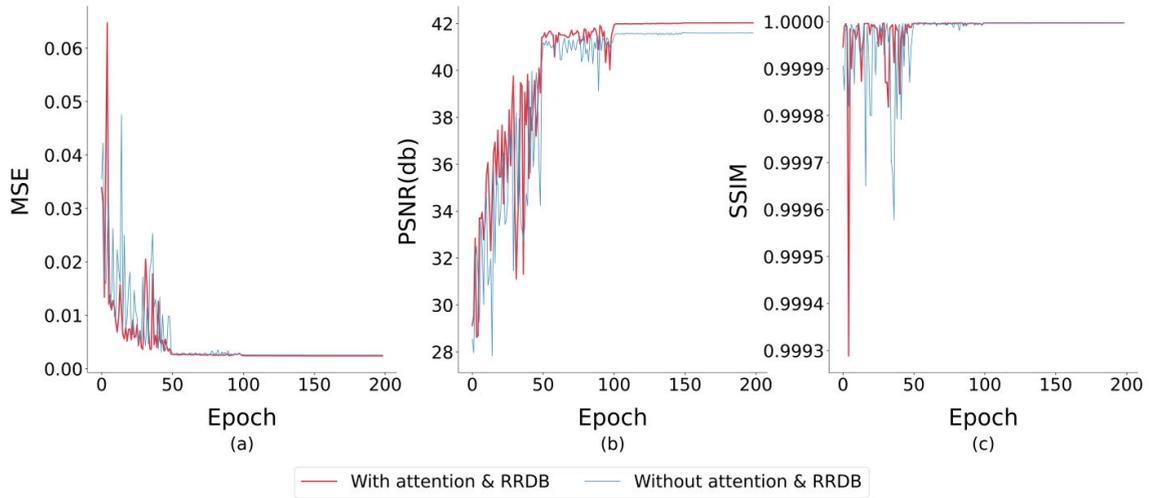
FIGURE 14. Differences produced by the DDSA model with and without Convolutional Block Attention Module (CBAM) and Residual in Residual dense blocks (RRDB) showing the *MSE, PSNR* and *SSIM vs.* the epoch. *Interpretive note: the MSE is relatively lower and PSNR is higher for the module with attention layer utilizing Residuals in the Residual Dense Block.*

FIGURE 15. Differences produced by the DDSA model with and without Convolutional Block Attention Module (CBAM), Residual in Residual dense blocks (RRDB) and Residual dense blocks (RDB) showing the *MSE, PSNR* and *SSIM vs.* the epoch. *Interpretive note: the MSE is relatively lower and PSNR is higher for the module with attention layer utilizing Residuals in the Residual Dense Block.*
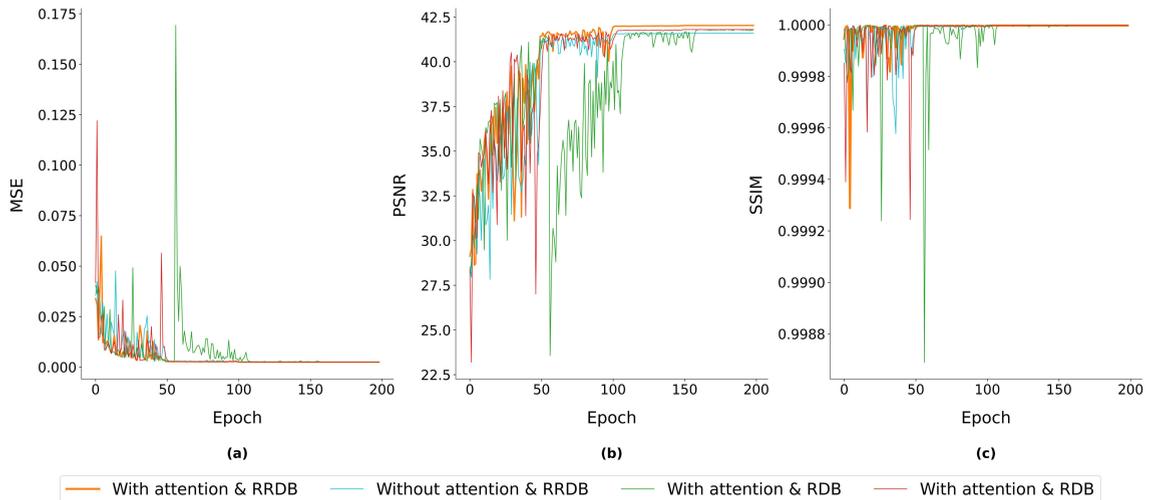
further improvement in the signal-to-noise ratio at approximately 100 epochs. The comparison highlights the contribution of CBAM in improving the focus on essential feature representation, which has delivered improved image quality both on visual and model performance benchmarks.

In Figure 14, we show the performance with addition of the CBAM to the RDB variation of the proposed DDSA model. The addition of attention layer seems to improve the performance to bring its accuracy to the second best model in terms of *PNSR* and as well as the other metrics. The introduction of RRDB instead of the RDB layer in subsequent iteration further improved the performance and therefore reinforced the theoretical and experimental evidence in literature. Overall the combination of CBAM and RRDB layers contributed to the DDSA model achieving state of the art performance within 100 epochs in contrast to benchmark

models, as noted earlier, despite pro-longed training under performed.

By contrast, the removal of both the CBAM and the RRDB layers while using standard RDB layers, as in Figure 15 demonstrates significantly uneven performance on lower epochs followed by stable performance well after 100 epochs. The addition of attention layer to the RDB model improves the performance to bring its accuracy to the second-best model in terms of *PNSR* and as well as the other metrics. The introduction of grouped RRDB instead of the RDB layer in subsequent iteration further improved the performance. Reinforcing the theoretical and experimental evidence in the literature, RRDB blocks mitigated the noise through the learning of hierarchical features, contributing to an enhanced quality output.

Overall the combination of CBAM and RRDB layers contributed to the DDSA model achieving state of the art

performance within 100 epochs in contrast to benchmark models, as noted earlier, which despite pro-longed training under performed.

VI. CONCLUSION, LIMITATIONS, AND FURTHER OUTLOOK

A. CONCLUSIONS

This study reported the development and evaluation of novel deep downscaling methods for surface albedo mapping at high spatial resolutions (i.e., 500 m, 333 m and 250 m), from low resolution (1 km) sky images, utilising image super-resolution models. In the context of solar energy gains and energy consumption estimations, we require granular scale surface albedo map including a number of other weather inputs for instantaneous solar energy monitoring or climatic inputs for long-term solar energy feasibility studies. However, the sparsity of current ground monitoring sites, poor spatial resolution of existing satellite data in their raw form and a lack of primary data or inadequacy of traditional parametric models that do not perform well at poor spatial resolutions, warrant new scientific methods such as deep learning algorithms to generate accurate spatial maps at granular scales currently not available. To address this issue, the present study has harnessed a deep learning-based image-super resolution algorithm, that is traditionally applied to enhance natural photography, and to further test the newly proposed model. Deep downscaling model with attention, or the DDSA model, has been constructed by integrating Depth wise Separable Convolution, Residual in Residual dense blocks (RRDB), Convolutional Block Attention Module (CBAM) and sub-pixel convolution to efficiently learn from remotely sensed low spatial resolution (1 km) black sky albedo images. Using the proposed method, we successfully produced high resolution representation of surface albedo maps, whose agreement with ground observations were aptly verified. Evaluated against benchmark models, and observations from ground validation sites, the proposed model demonstrated excellent performance advantage in terms of producing the surface albedo maps at high resolutions. The quantitative performance of standard models such as SRCNN, EDSR, ESPCN and RDN models lagged behind that of the proposed DDSA model. The main findings of the work are as follows:

- In terms of qualitative evaluations, the DDSA model achieved the closest visual representation of ground truth image based on model predictions for the OzFlux ground stations.
- A 16-quantile intensity analysis of a portion of ground truth pixel distribution of model predictions across all models reveal the best reproduction of ground truth by the proposed model with no perceptible artefacts, whereas the RDN model generating the closest second match with grid artefacts.
- Error maps showing the visual distribution of the magnitude of errors between the ground truth and model predictions showed that the proposed model predictions

were closer to ground truth compared to the benchmark models

- Contrasting the predictive capability of the DDSA using deep downscaled 500 m, 333 m and 250 m predictions against the OzFlux station observation and PROBA-V/SPOT, we noted that the proposed model registered lower deviations from ground truth for Cumberland, Tumbarumba, Litchfield and Wombat Creek having homogenous land surface classes evergreen broadleaf and woody savannas. Similarly, the deep downscaled predictions for 500 m surface albedo showed much lower deviation from the ground truth than surface albedo estimates from MODIS 500 m surface albedo product for Cumberland and Calperum sites with homogenous surface characteristics. The deviations showed no clear trends for Cape Tribulation with a mixed foreest surface and Wombat Stringbark with Evergreen Broadleaf vegetation.
- In spite of the superior performance elsewhere, the proposed model under performed in its capability when compared with the OzFlux station observations for Calperum and Cape Tribulation locations with heterogeneous land surface of mixed forest and the latter closed shrub lands.
- Ablation studies performed on DDSA revealed that the combination of convolutional Block Attention Module and the Residual in Residual dense blocks enhanced the overall model performance without significant impacts on the overall computational efficiency.

Although this study has validated the method for granular scale prediction of 10-day black sky surface albedo in Australian mainland in the summer period, the proposed DDSA model could potentially be applied to other locations and seasons to expand its feasibility more widely. In spite of the excellent performance obtained, we also note that the contributions of the attention mechanisms in the proposed DDSA model can be improved with a combination of the RRDB and other attention mechanism methods. In terms of balancing the accuracy and the computational efficiency, especially when using large kernel sizes such as Extremely Separated Convolutions (XSeptConv) [60], these networks could be further researched as a replacement for depth wise convolution layer in a future study.

B. LIMITATIONS, FURTHER OUTLOOK, AND FUTURE RESEARCH

The limitations in satellite sensor capabilities and sparsity of ground-based measurement sensors warrants the need for weather and climate observations to be generated at higher granularity, especially in the context of applications that require higher spatial precision. By demonstrating the state-of-the-art downscaling capabilities, this work has presented the proposed DDSA model as excellent scientific utility with advanced predictive skills compared to the benchmark models e.g., SRCNN, EDSR, EPSCN, and RDN for

generating granular scale albedo observations across time and space. Assessed against ground observations at six FluxNet locations, MODIS surface alebdo estimations, and contrasted against PROBA-V/SPOT observations, the DDSA model reaffirmed itself as a promising approach to address the challenge of sparsity in surface albedo observations for critical applications in solar PV deployment and urban planning applications. The ability to generate high resolution albedo observations is expected to aid renewable energy industries in modelling systems such as the Agri-PV, which combines food production (in agricultural space) and energy production, allowing land to be used for agricultural production and also in the generation of solar power. In particular, AgriPV scenarios and transition pathways from the mono facial to the bifacial solar panels can boost the contributions of solar renewable energies in the energy mix, while giving rise to innovative energy monitoring models. Apart from the immediate application in Agri-PV area, the proposed DDSA model could also be applied in the context of energy planning strategies to address the negative effects of surface albedo, especially in urban settings which is known to contribute to urban climate phenomena like the Urban Heat Island (UHI) effect [61]. Especially downscaled surface alebdo predictions could be combined with equivalent remote sensing platforms such as MODIS to generate more precise predictions through ensembling approaches.

In spite of the success of the objective model, its predictive skills are somewhat limited in terms of the low temporal frequency, spatial resolution and available spectral frequencies in the dataset. Future research which harnesses the image datasets with a higher temporal frequency and better spatial resolution using sources such as Himawari 8/9 satellites, or utilising hyper-spectral datasets, may improve the model skill in terms of better spatial and temporal granularity, while enhancing the robustness of the predictions. Furthermore, since the scope of this study was limited to some parts of mainland Australia, limiting the range of surface types and constrained by the availability of ground validation stations, different geographical regions and landscapes could pose challenges in the application of the proposed DDSA model. Given the nexus between surface types and surface albedo, future research in evaluating the role of land cover inputs in enhancing the generalisability of the objective model across geographical regions in addition to spectral inputs is under investigation.

The lack of availability of ground observations over better temporal frequencies and length has limited the evaluation of the predictive ability of the proposed model, in the context of extreme or unseasonal weather conditions. In the few instances where data were available, the model predictions indicated higher deviations from the ground truth. For instance, deviations between ground observations and predictions were higher for FluxNet location Cumberland (see Figure 10a) on 24 January 2016. A review of the weather conditions preceding immediately before these observations, as provided by the nearest Australian Bureau of Meteorology

Station Richmond, RAAF indicates that the 4 days of temperatures were exceeded 30°C, of which the 3 days of temperatures were greater than 35°C with significant variability. This period was immediately followed up by a period of cooler weather with average temperature of 27°C over the next observation period. Similarly, FluxNet location Cumberland (Figure 10a) also experience a high degree of weather variability in the period leading up to the observations taken on 13 January 2016 with temperatures exceeding 40°C in the 3 days preceding the observation. The period which followed also experienced a high degree of variability. These findings indicate that the objective model suffers a decline in predictive skill in conditions of climate extremes and in the presence of heterogeneous vegetation cover. requiring further investigation. To address performance declines in climate extremes the model could potentially be enhanced through the pairing of remotely sensed image datasets with additional climate variables such as air temperature, rainfall and wind speed data from ground stations. However, this too poses a challenge, given the sparsity of weather stations that could provide spatially congruent weather observations.

It is noteworthy that the contribution of the attention mechanism within the proposed DDSA model was significant in respect to improved predictions at higher spatial resolutions. However, this aspect may be further improved with a better combination of the RRDB and the other types of attention mechanisms. In terms of balancing the accuracy and the computational efficiency, especially when using large kernel sizes, alternative approaches such as the Extremely Separated Convolutions (XSeptConv) [60] could be further explored as a replacement for the depth wise convolutions used in the present model.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT
Sagthitharan Karalasingham: conceptualization, visualization, investigation, methodology, software, data curation, writing—original draft. Ravinesh C. Deo: conceptualization, visualization, investigation, supervision, writing—review and editing. David Casillas-Pérez: writing—review and editing. Nawin Raj: supervision, writing—review and editing. Sancho Salcedo-Sanz: supervision, writing—review and & editing.

DECLARATION OF COMPETING INTEREST
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGMENT

VII. LIST OF ACRONYMS

Table 2 shows the list of acronyms and metrics used in this paper.

TABLE 2. List of acronyms.

Acronym	Full name
AI	Artificial Intelligence
ANN	Artificial Neural Networks
BCSD	Bias Corrected Spatial Disaggregation
CBAM	Convolutional Block Attention Module
CNN	Convolutional Neural Networks
COMS	Communication, Ocean and Meteorological Satellite
DCNN	Deep Convolutional Neural Networks
DDSA	Deep Downscaling Spectral model with Attention
DSC	Depth Separable Convolution
ECV	Essential Climate Variables
EDSR	Enhanced Deep Super-Resolution network
ESM	Earth System Model
ESPCN	Efficient Sub-Pixel Convolutional Neural Network
GBOV	Ground-Based Observations for Validation
GCOS	Global Climate Observing System
GDAL	Geospatial Data Abstraction Library
GPU	Graphical Processing Units
HR	High Resolution
HRRR	High-Resolution Rapid Refresh
LR	Low Resolution
MSE	Mean Square Error
MSSL	Multitask Sparse Structure Learning
PSNR	Signal-to-Noise-Ratio
PV	Photo Voltaic
RDB	Residual Dense Blocks
RDN	Residual Dense Networks
RRDB	Residual in Residual Dense Blocks
SDSM	Statistical Downscaling
SRCNN	Super-Resolution Convolution Neural Network
SSIM	Structural Similarity Index
SVM	State Vector Machines

REFERENCES

[1] S. Bouckaert, A. F. Pales, C. McGlade, U. Remme, B. Wanner, L. Varro, D. D'Ambrosio, and T. Spencer, "Net zero by 2050: A roadmap for the global energy sector," OECD Publishing, Paris, France, Tech. Rep., 2021. [Online]. Available: https://www.oecd-ilibrary.org/energy/net-zero-by-2050_c8328405-en

[2] M. Smid and A. C. Costa, "Climate projections and downscaling techniques: A discussion for impact studies in urban systems," Int. J. Urban Sci., vol. 22, no. 3, pp. 277–307, Nov. 2017.

[3] Climate Change and Cities, International Conference on Climate Change and Cities, UN-Habitat UNHSP, Nairobi, Kenya, 2018.

[4] P. R. Shukla et al., "Climate change and land: An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems," IPCC, Tech. Rep., 2019. [Online]. Available: https://www.ipcc.ch/srccl/cite-report/

[5] H. Ziar, F. F. Sönmez, O. Isabella, and M. Zeman, "A comprehensive albedo model for solar energy applications: Geometric spectral albedo," Appl. Energy, vol. 255, Dec. 2019, Art. no. 113867.

[6] F. Cotana, F. Rossi, M. Filipponi, V. Coccia, A. L. Pisello, E. Bonamente, A. Petrozzi, and G. Cavalaglio, "Albedo control as an effective strategy to tackle global warming: A case study," Appl. Energy, vol. 130, pp. 641–647, Oct. 2014.

[7] X. Sun, M. R. Khan, C. Deline, and M. A. Alam, "Optimization and performance of bifacial solar modules: A global perspective," Appl. Energy, vol. 212, pp. 1601–1610, Feb. 2018.

[8] R. Kopecek and J. Libal, "Bifacial photovoltaics 2021: Status, opportunities and challenges," Energies, vol. 14, no. 8, p. 2076, Apr. 2021.

[9] M. Khenar, S. Taheri, A.-M. Cretu, S. Hosseini, and E. Pouresmaeil, "PSO-based modeling and analysis of electrical characteristics of photovoltaic module under nonuniform snow patterns," IEEE Access, vol. 8, pp. 197484–197498, 2020.

[10] I. E. Agency. (2020). Energy Efficient Buildings in Armenia: A Roadmap. [Online]. Available: https://www.iea.org/reports/energy-efficient-buildings-in-armenia-a-roadmap

[11] G. M. Tina, F. B. Scavo, L. Merlo, and F. Bizzarri, "Comparative analysis of monofacial and bifacial photovoltaic modules for floating power plants," Appl. Energy, vol. 281, Jan. 2021, Art. no. 116084.

[12] N. Riedel-Lyngskær, M. Ribaconka, M. Pó, A. Thorseth, S. Thorsteinsson, C. Dam-Hansen, and M. L. Jakobsen, "The effect of spectral albedo in bifacial photovoltaic performance," Sol. Energy, vol. 231, pp. 921–935, Jan. 2022.

[13] G. C. O. System. (2022). Albedo, Essential Climate Variable (ECV) Factsheet. [Online]. Available: https://ane4bf-datap1.s3.eu-west-1.amazonaws.com/wmod8_climatedata/s3fs-public/albedo_ecv_factsheet_201905.pdf

[14] U. O. M. B. Professor Crystal Schaaf's Lab. (2022). MOD43B3 Albedo Product. [Online]. Available: https://www.umb.edu/spectralmass/terra_aqua_modis/modis_user_guide_v004/mod43b3_albedo_product#:~:text=The%20MOD43B3%20Albedo%20Product%20(MODIS,%2C%20and%200.3%2D5.0%C2%B5m)

[15] J. Stein, C. Reise, J. B. Castro, G. Friesen, G. Maugeri, E. Urrejola, and S. Ranta, "Bifacial photovoltaic modules and systems: Experience and results from international research and pilot applications," Sandia Nat. Lab. (SNL-NM), Albuquerque, NM, USA, Tech. Rep. SAND-2021-4835R; IEA-PVPS T13-14:2021 695675, 2021.

[16] B. Marion, "Measured and satellite-derived albedo data for estimating bifacial photovoltaic system performance," Sol. Energy, vol. 215, pp. 321–327, Feb. 2021.

[17] J. B. Jahangir, M. Al-Mahmud, M. S. S. Shakir, A. Haque, M. A. Alam, and M. R. Khan, "A critical analysis of bifacial solar farm configurations: Theory and experiments," IEEE Access, vol. 10, pp. 47726–47740, 2022.

[18] T. Vandal, E. Kodra, and A. R. Ganguly, "Intercomparison of machine learning methods for statistical downscaling: The case of daily and extreme precipitation," Theor. Appl. Climatol., vol. 137, nos. 1–2, pp. 557–570, Jul. 2019.

[19] M. S. Khan, P. Coulibaly, and Y. Dibike, "Uncertainty analysis of statistical downscaling methods using Canadian global climate model predictors," Hydrolog. Processes, vol. 20, no. 14, pp. 3085–3104, 2006.

[20] H. Su, Z. Xiong, X. Yan, and X. Dai, "An evaluation of two statistical downscaling models for downscaling monthly precipitation in the Heihe river basin of China," Theor. Appl. Climatol., vol. 138, nos. 3–4, pp. 1913–1923, Nov. 2019.

[21] A. H. Nury, A. Sharma, L. Marshall, and R. Mehrotra, "Characterising uncertainty in precipitation downscaling using a Bayesian approach," Adv. Water Resour., vol. 129, pp. 189–197, Jul. 2019.

[22] S. Agrawal, L. Barrington, C. Bromberg, J. Burge, C. Gazen, and J. Hickey, "Machine learning for precipitation nowcasting from radar images," 2019, arXiv:1912.12132.

[23] K. Tsipis, "3Q: Machine learning and climate modeling," MIT News, Cambridge, MA, USA, Tech. Rep., 2019.

[24] J.-M. Yeom, R. C. Deo, J. F. Adamwoski, T. Chae, D.-S. Kim, K.-S. Han, and D.-Y. Kim, "Exploring solar and wind energy resources in North Korea with COMS MI geostationary satellite data coupled with numerical weather prediction reanalysis variables," Renew. Sustain. Energy Rev., vol. 119, Mar. 2020, Art. no. 109570.

[25] J.-M. Yeom, R. C. Deo, J. F. Adamowski, S. Park, and C.-S. Lee, "Spatial mapping of short-term solar radiation prediction incorporating geostationary satellite images coupled with deep convolutional LSTM networks for South Korea," Environ. Res. Lett., vol. 15, no. 9, Sep. 2020, Art. no. 094025.

[26] D. Guijo-Rubio, A. M. Durán-Rosal, P. A. Gutiérrez, A. M. Gómez-Orellana, C. Casanova-Mateo, J. Sanz-Justo, S. Salcedo-Sanz, and C. Hervás-Martínez, "Evolutionary artificial neural networks for accurate solar radiation prediction," *Energy*, vol. 210, Nov. 2020, Art. no. 118374.

[27] L. Cornejo-Bueno, C. Casanova-Mateo, J. Sanz-Justo, and S. Salcedo-Sanz, "Machine learning regressors for solar radiation estimation from satellite data," *Sol. Energy*, vol. 183, pp. 768–775, May 2019.

[28] (Dec. 2016). *Sentinels Down Under*. [Online]. Available: https://sentinels.copernicus.eu/web/sentinel/news//asset_publisher/xR9e/content/sentinels-down-under

[29] M. Claverie, J. Ju, J. G. Masek, J. L. Dungan, E. F. Vermote, J.-C. Roger, S. V. Skakun, and C. Justice, "The harmonized Landsat and Sentinel-2 surface reflectance data set," *Remote Sens. Environ.*, vol. 219, pp. 145–161, Dec. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425718304139

[30] (Jun. 2020). *Copernicus Global Land Service—Surface Albedo*. [Online]. Available: https://land.copernicus.eu/global/products/sa?qt-sa_characteristics=4

[31] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2015.

[32] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1646–1654.

[33] Z. Wang, J. Chen, and S. Hoi, "Deep learning for image super-resolution: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 10, pp. 3365–3387, Mar. 2020.

[34] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1874–1883.

[35] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2472–2481.

[36] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," in *The Handbook of Brain Theory and Neural Networks*, vol. 3361, no. 10. Cambridge, MA, USA: MIT Press, 1995, p. 1995.

[37] L. Dang, P. Pang, and J. Lee, "Depth-wise separable convolution neural network with residual connection for hyperspectral image classification," *Remote Sens.*, vol. 12, no. 20, p. 3408, Oct. 2020.

[38] L. Kaiser, A. N. Gomez, and F. Chollet, "Depthwise separable convolutions for neural machine translation," 2017, *arXiv:1706.03059*.

[39] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5353–5360.

[40] Y. Guo, Y. Li, L. Wang, and T. Rosing, "Depthwise convolution is all you need for learning multiple visual domains," in *Proc. AAAI Conf. Artif. Intell. (AAAI)*, vol. 33, 2019, pp. 8368–8375.

[41] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1251–1258.

[42] C. Runkel, C. Etmann, M. Möller, and C.-B. Schönlieb, "Depthwise separable convolutions allow for fast and memory-efficient spectral normalization," 2021, *arXiv:2102.06496*.

[43] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 391–407.

[44] D.-W. Kim, J. R. Chung, and S.-W. Jung, "GRDN: Grouped residual dense network for real image denoising and GAN-based real-world noise modeling," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2019, pp. 1–9.

[45] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. C. Loy, "ESRGAN: Enhanced super-resolution generative adversarial networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV) workshops*, 2018, pp. 1–16.

[46] M. Carrasco, "How visual spatial attention alters perception," *Cognit. Process.*, vol. 19, no. S1, pp. 77–88, Sep. 2018.

[47] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3156–3164.

[48] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19.

[49] X. Zhao, J. Zhang, J. Tian, L. Zhuo, and J. Zhang, "Residual dense network based on channel-spatial attention for the scene classification of a high-resolution remote sensing image," *Remote Sens.*, vol. 12, no. 11, p. 1887, Jun. 2020.

[50] A. Aitken, C. Ledig, L. Theis, J. Caballero, Z. Wang, and W. Shi, "Checkerboard artifact free sub-pixel convolution: A note on sub-pixel convolution, resize convolution and convolution resize," 2017, *arXiv:1707.02937*.

[51] EU. (2022). *Copernicus Global Land Service*. [Online]. Available: https://land.copernicus.eu/global/products/sa

[52] E. C. J. R. Centre. (2022). *Ground-Based Observations for Validation (GBOV) of Copernicus Global Land Products*. [Online]. Available: https://gbov.acri.fr/

[53] T. E. R. Network. (2022). *Australia and New Zealand Flux Research and Monitoring*. [Online]. Available: https://www.ozflux.org.au/

[54] R. Song, S. Kharbouche, and J.-P. Müller, "Ground-based observations for validation (GBOV) of Copernicus global land products," in *Proc. Geophys. Res. Abstr.*, vol. 21, 2019, p. 17608.

[55] K. Wang, S. Liang, C. L. Schaaf, and A. H. Strahler, "Evaluation of moderate resolution imaging spectroradiometer land surface visible and shortwave albedo products at FLUXNET sites," *J. Geophys. Res.*, vol. 115, no. D17, pp. 1–8, 2010.

[56] S. Gopinathan and M. S. Gayathri, "A study on image enhancement techniques using YCbCr color space methods," *Int. J. Adv. Eng. Res. Sci.*, vol. 3, no. 8, 2016, Art. no. 236818.

[57] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[58] A. R. Reibman, R. M. Bell, and S. Gray, "Quality assessment for super-resolution image enhancement," in *Proc. Int. Conf. Image Process.*, Oct. 2006, pp. 2017–2020.

[59] M. Deudon, A. Kalaitzis, I. Goytom, M. R. Arefin, Z. Lin, K. Sankaran, V. Michalski, S. E. Kahou, J. Cornebise, and Y. Bengio, "HighRes-Net: Recursive fusion for multi-frame super-resolution of satellite imagery," 2020, *arXiv:2002.06460*.

[60] J. Chen, Z. Lu, J.-H. Xue, and Q. Liao, "XSepConv: Extremely separated convolution," 2020, *arXiv:2002.12046*.

[61] E. Morini, A. Touchaei, B. Castellani, F. Rossi, and F. Cotana, "The impact of albedo increase to mitigate the urban heat island in Terni (Italy) using the WRF model," *Sustainability*, vol. 8, no. 10, p. 999, Oct. 2016.

SAGTHITHARAN KARALASINGHAM (Graduate Student Member, IEEE) received the M.Sc. degree (Hons.) in mathematics and statistics from the University of Southern Queensland (UniSQ), Australia, in 2020. He completed an Australian Postgraduate Research Internship at Senetas Corporation supported by the Defence Science Institute. Prior to academia, he was in technology startup companies in Singapore and Australia. He also worked in innovation and economic policy and technology consulting. He is currently undertaking research in computer vision and deep learning for remote sensing energy demand patterns under a UniSQ Ph.D. Scholarship. His current research interests include science with artificial intelligence (AI), bio-inspired computer vision, human–centered AI, and AI-driven remote sensing. He was a recipient of the UniSQ ReDTrain Summer Research Scholarship and the Bertelsmann Scholarship Challenge (AI Track) Finalist.

RAVINESH C. DEO (Senior Member, IEEE) leads the UniSQ's Advanced Data Analytics Research Laboratory as a Professor at the University of Southern Queensland (UniSQ), Australia. His expertise is in artificial intelligence and machine learning methods for renewable energy and climate science. He is among scientists and social scientists who have demonstrated significant broad influence, reflected in the publication of multiple papers frequently cited by their peers. He also leads cross-disciplinary research in deep learning and artificial intelligence having supervised more than 30 Ph.D./M.Sc. degrees. He has published more than 270 articles, and seven books with an H-index of 60 that has cumulative citations that exceeds 11 600. His publications rank in top 1% by citations for field and publication year in the Web of Science citation index. He was a recipient of the Employee Excellence Awards, the Elsevier Highly Cited Paper Awards, and the Publication Excellence and Teaching Commendations. He is a 2021 Clarivate Analytics Highly Cited Author.

DAVID CASILLAS-PÉREZ received the B.S. degree in telecommunication engineering and the M.S. and Ph.D. degrees in electronic control systems from the Universidad de Alcalá, Spain, in 2013, 2014, and 2019, respectively. He is currently an Assistant Professor with the Department of Signal Processing and Communications, Universidad Rey Juan Carlos, Madrid, Spain. His research interests include the development of machine learning algorithms with applications in different fields, such as computer vision, mobile communication systems, and renewable energy systems.

NAWIN RAJ received the B.Sc., B.Ed., PGDMA, and M.Sc. degrees in computational fluid dynamics from The University of the South Pacific and the Ph.D. degree from the School of Mathematics, Physics and Computing, University of Southern Queensland (UniSQ), Queensland, Australia, in 2015. From 2007 to 2010, he was a Lecturer at the Fiji National University. From 2015 to 2016, he worked as a Mathematics Learning Advisor at UniSQ, where he is currently a Senior Lecturer. He is also a member of the UniSQ's Advanced Data Analytics Research Laboratory, UniSQ. His research interests include artificial intelligence, deep learning, non-linear oscillation, computational fluid dynamics, and oceanography.

SANCHO SALCEDO-SANZ was born in Madrid, Spain, in 1974. He received the B.S. degree in physics from the Universidad Complutense de Madrid, Spain, in 1998, the Ph.D. degree in telecommunications engineering from the Universidad Carlos III de Madrid, Spain, in 2002, and the Ph.D. degree in physics from the Universidad Complutense de Madrid, in 2019. He spent one year at the School of Computer Science, University of Birmingham, U.K, as a Postdoctoral Research Fellow. He is currently a Full Professor at the Department of Signal Processing and Communications, Universidad de Alcalá, Spain. He has coauthored more than 220 international journal articles in the field of machine learning and soft-computing and its applications. His research interests include soft-computing techniques, hybrid algorithms, and neural networks in different problems of science and technology.

• • •