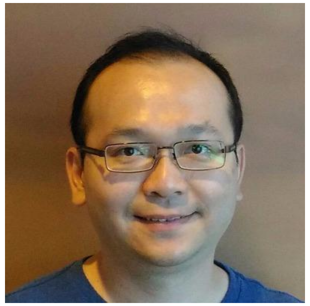
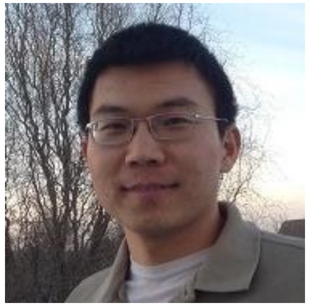
IBM Research

Image-Based Soil Organic Carbon Estimation from Multispectral Satellite Images with Fourier Neural Operator and Structural Similarity

Ken C. L. Wong[1], Levente Klein[2], Ademir Ferreira da Silva[3], Hongzhi Wang[1], Jitendra Singh[4], Tanveer Syeda-Mahmood[1]

[1]IBM Research Almaden, San Jose, CA, USA
[2]IBM Research, Yorktown Heights, NY, USA
[3]IBM Research, Rio De Janeiro, RJ, Brazil
[4]IBM Research, Gurgaon, HR, India

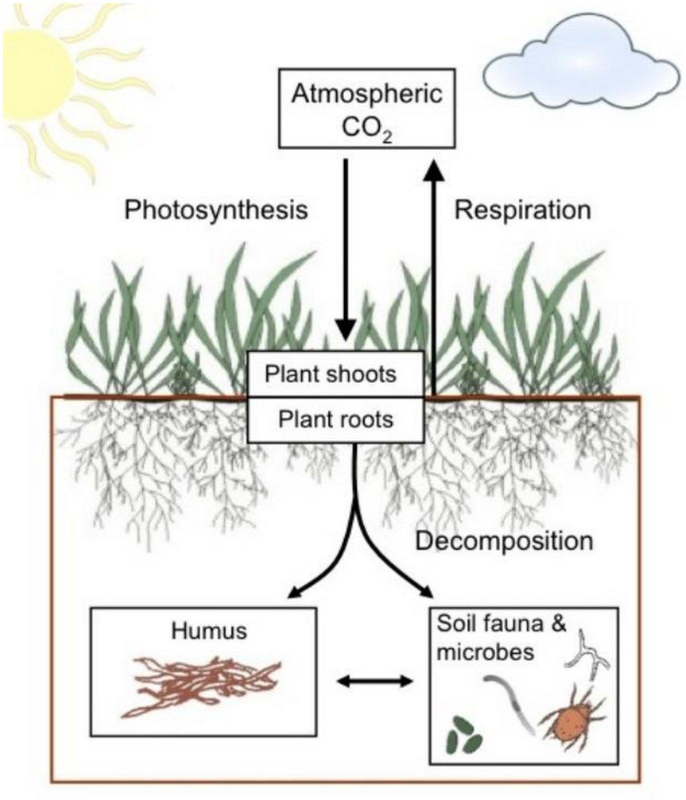


Soil Organic Carbon (SOC) Estimation

- Soil organic carbon sequestration
  - The process of capturing and storing atmospheric CO_2 to soil through natural carbon cycle (e.g., plants)
  - Soils can sequester 1.85 petagrams (10^{15} grams) of carbon per year – important for climate change mitigation
  - Effectiveness can be improved by proper land use – monitoring and managing SOC level are important

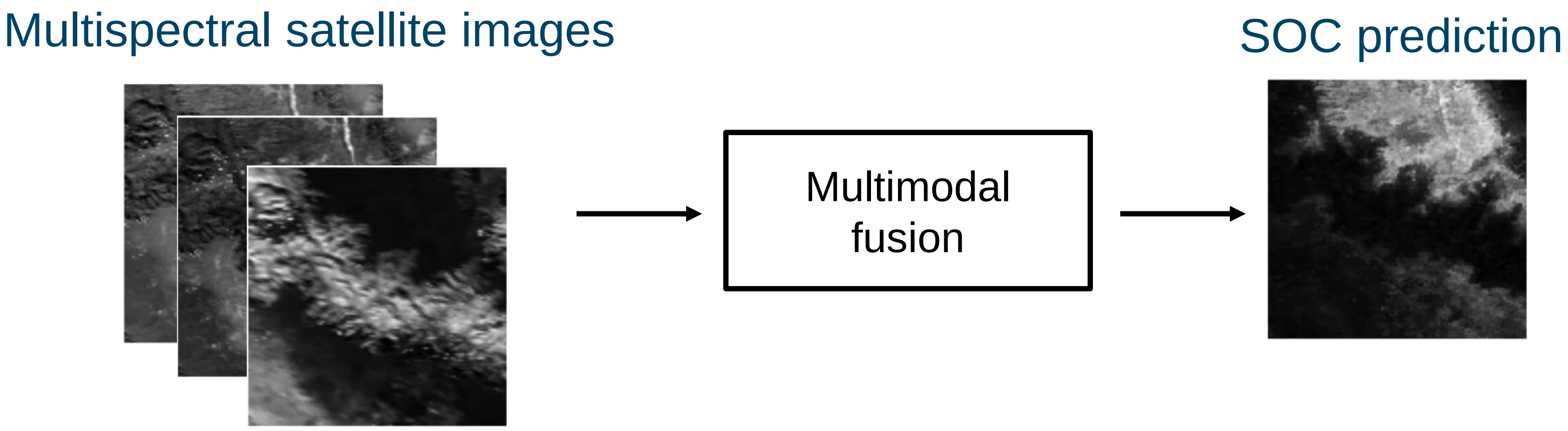
- SOC estimation
  - Important for monitoring SOC level
  - Traditional soil sampling and lab tests – impractical at a global scale
  - Estimation at a global scale can be achieved by satellite imaging



Todd A. Ontl and Lisa A. Schulte. "Soil carbon storage." *Nature Education Knowledge* 3, no. 10 (2012).

Data

- Data
  - Satellite data of MODIS (Moderate Resolution Imaging Spectroradiometer) from two satellites
    - Six spectral bands: blue, green, red, near infrared, and shortwave infrared (2 bands)
    - Provide information such as water, vegetation, soil moistures – related to SOC level
  - Organic carbon data from SoilGrids (https://www.isric.org/explore/soilgrids)
    - Global gridded soil information
  - Goal: multimodal fusion of satellite data for SOC remote sensing



Fourier Neural Operator

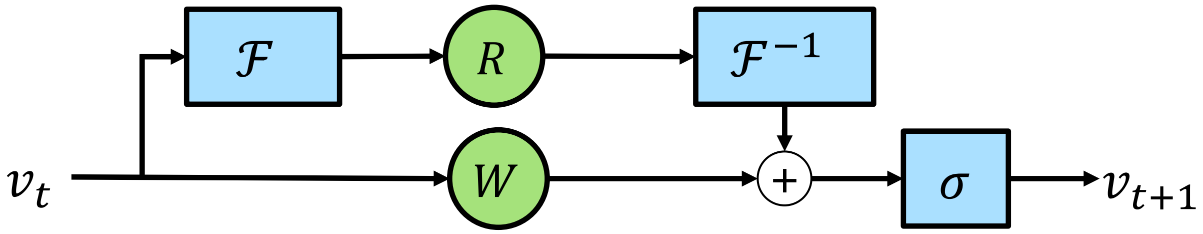
- Fourier neural operator (FNO) Li, Zongyi, et al. "Fourier neural operator for parametric partial differential equations." *ICLR* (2021)
  - Originally proposed to learn function mappings in PDEs (e.g., Navier-Stokes equations)
  - Formulated in the continuous space based on the idea of Green's function
  - Neural operator: iterative updates

Analogue to Green's function

$$v_{t+1}(x) := \sigma\left(W v_t(x) + (\mathcal{K} v_t)(x)\right) \quad \text{with} \quad (\mathcal{K} v_t)(x) := \int_D \kappa(x-y) v_t(y)\, dy, \quad \forall x \in D$$

Learnable weights in Fourier domain

  - FNO – efficient computation with Fourier transform: $(\mathcal{K} v_t)(x) = \mathcal{F}^{-1}\left(R \cdot (\mathcal{F} v_t)\right)(x), \quad \forall x \in D$

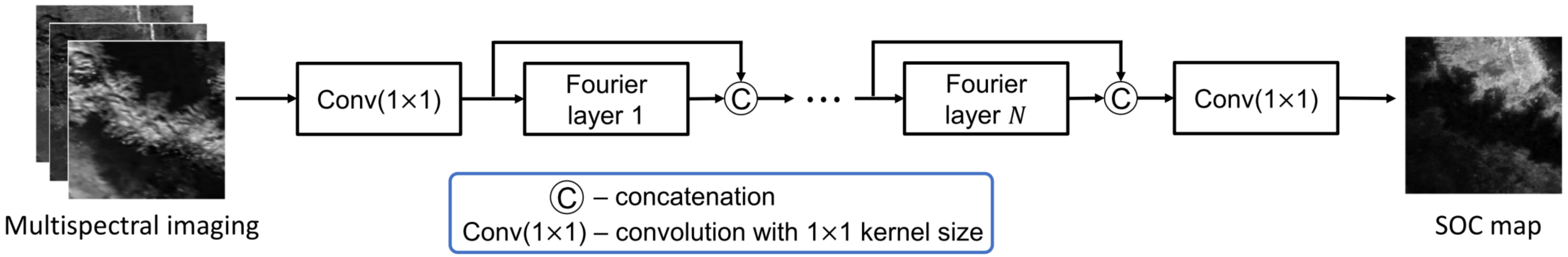


Fourier layer

FNO-DenseNet

- FNO
  - ✓ Intrinsic zero-shot super-resolution, global receptive field
  - × Large number of parameters – different weight matrices (R) at different locations in the Fourier domain
    - Number of weights proportional to image size

- Our proposed FNO-DenseNet
  - Use shared weights in the Fourier domain – reduce numbers of parameters by hundreds of times
  - Use the DenseNet idea to improve convergence and accuracy



FNO-DenseNet

Loss Functions

- Loss functions
  - Pixel-based loss functions (e.g., MAE, MSE) did not work well in experiments
  - Structural similarity (∈[-1, 1]) – measures difference in structural information
    - Compares the local luminance, contrast, and structures between two images
    - Considers interdependencies among local pixels (window) – good for textural analysis

$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = \frac{1}{M}\sum_{j=1}^{M}\frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

  - DSSIM (Dissimilarity) loss for minimization

$$\mathrm{DSSIM} = 0.5 \times (1 - \mathrm{SSIM}) \in [0, 1]$$

  - Weighted sum of MAE and DSSIM loss – provides the overall best results

$$w \times \mathrm{MAE} + \mathrm{DSSIM}$$

Wang, Zhou, et.al. "Image quality assessment: from error visibility to structural similarity." *IEEE Transactions on Image Processing* 13, no. 4 (2004)

Experiments

- Tested four approaches
  - Random forest – 10 trees, max depth of 10
  - Modified V-Net (3 encoding blocks, 3 decoding blocks, initial filters = 16, params = 342K)
  - Original FNO (8 Fourier layers, 32 filters per layer, params = 34M)
  - Proposed FNO-DenseNet (8 Fourier layer, growth rate = 24 filters, params = 64K) – 500 times fewer params

- Data split (3059 samples)
  - Training – 50%
  - Validation – 20%
  - Testing – 30%

- Image augmentation
  - Rotation – 30 degrees
  - Shifting – 20%
  - Zooming – [0.8, 1.2]
  - Random flip
  - 80% chance to be transformed

Results

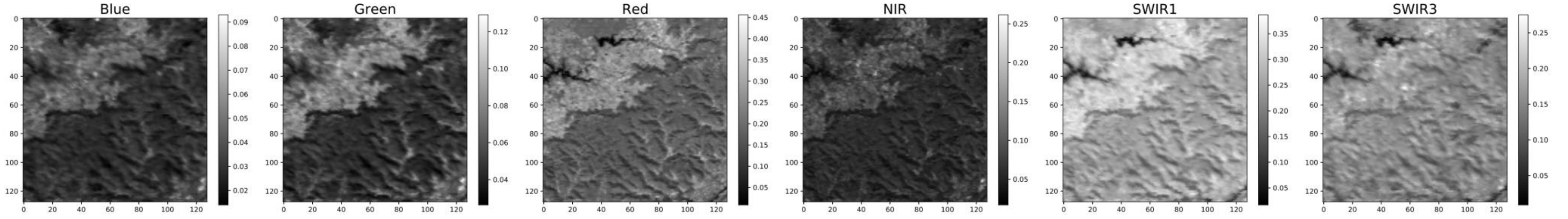
- Metrics
  - RMSE – root mean squared error (g/kg)
  - MAPE – mean absolute percentage error (%)
  - SSIM – structural similarity (∈[-1, 1], 1 is the best)

- Observations
  - FNO-DenseNet – best results with smallest model size (64K params)
  - Random forest – worst results
  - MAE-only loss – good pixel-wise accuracy, bad structural similarity
  - DSSIM-only loss – good structural similarity, bad pixel-wise accuracy
  - MAE + DSSIM loss – good structural similarity and good pixel-wise accuracy

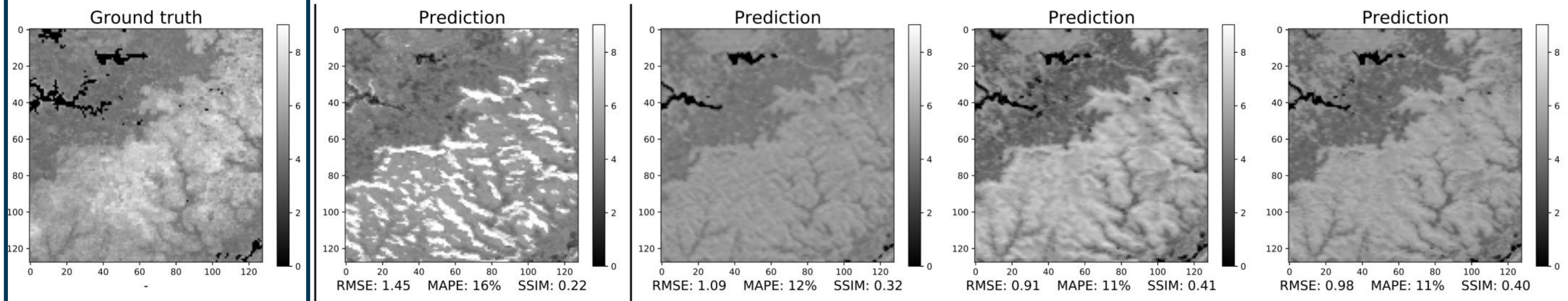
Loss	MAE			DSSIM			MAE + DSSIM		
Metric	RMSE	MAPE	SSIM	RMSE	MAPE	SSIM	RMSE	MAPE	SSIM
Random forest	2.50±2.37	45.40±46.30	0.07±0.08	——	——	——	——	——	——
Modified V-Net	2.12±2.26	30.12±23.00	0.07±0.06	2.21±2.25	33.44±22.23	**0.20±0.13**	2.00±1.99	29.79±21.11	0.17±0.12
FNO	**1.97±2.03**	28.08±21.13	0.10±0.09	2.31±2.34	35.57±24.22	**0.20±0.13**	1.96±1.96	28.32±18.29	**0.18±0.13**
FNO-DenseNet	1.98±2.09	**27.24±19.25**	**0.11±0.10**	**2.16±2.13**	**32.33±19.48**	**0.20±0.13**	**1.89±1.75**	**27.75±18.12**	**0.18±0.12**
		✓	✗		✗	✓		✓	✓

Visual Comparison

Multispectral inputs



SOC outputs



Ground truth | Random forest (MAE) | FNO-D (MAE) | FNO-D (DSSIM) | FNO-D (MAE + DSSIM)

Thanks!

clwong@us.ibm.com