JGR | Machine Learning and Computation

RESEARCH ARTICLE
10.1029/2024JH000174

# Sub-Pixel Displacement Estimation With Deep Learning: Application to Optical Satellite Images Containing Sharp Displacements

**Tristan Montagnon[1], Sophie Giffard-Roisin[1], Mauro Dalla Mura[2,3], Mathilde Marchandon[1,4], Erwan Pathier[1], and James Hollingsworth[1]**

[1]University Grenoble Alpes, University Savoie Mont Blanc, CNRS, IRD, University Gustave Eiffel, ISTerre, Grenoble, France, [2]University Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France, [3]Institut Universitaire de France (IUF), Paris, France, [4]Department of Earth and Environmental Sciences, Ludwig-Maximillians-Universität München, München, Germany

**Abstract** Optical image correlation is a powerful method for remotely constraining ground movement from optical satellite imagery related to natural disasters (e.g., earthquakes, volcanoes, landslides). This approach enables the characterization, and identification of the causal factors and mechanisms underlying such processes. By employing sub-pixel correlation algorithms, one can obtain highly accurate (m-to-cm level) displacement fields at high spatial resolution (dm-to-cm) by comparing satellite images acquired before and after a period of movement. However, this method generally assumes a homogeneous translation of all pixels within a given correlation window, which will lead to biased estimates of ground displacement if the real case is not well represented by such a simplification, especially when resolving ground displacements next to sharp gradients in displacement, such as those found in the near-field of earthquake surface ruptures. In this paper, we present an innovative deep learning method estimating sub-pixel displacement maps from optical satellite images for the retrieval of ground displacement. From the generation of a realistic simulated database, comprising Landsat-8 satellite image pairs containing simulated sub-pixel shifts and sharp discontinuities, we develop a Convolutional Neural Network able to retrieve sub-pixel displacements. The comparison to state-of-the-art correlation methods shows that our pipeline significantly reduces by 32% the estimation bias around fault ruptures, leading to more accurate characterization of the near-field strain in surface rupturing earthquakes. Application of our model to the 2019 Ridgecrest earthquake demonstrates the ability of our model to accurately and quickly resolve ground displacement using real satellite images. Code is made available at https://gricad-gitlab.univ-grenoble-alpes.fr/montagtr/cnn4l-discontinuities.

**Plain Language Summary** The precise estimation of ground displacement caused by natural hazards, such as earthquakes, volcanoes, landslides, as well as monitoring of glaciers, can be performed by comparing two optical satellite images of the same region acquired on different dates. The challenge resides in the fact that the ground motion is generally smaller than the satellite image resolution: sub-pixel precision is therefore critical. One solution, at the core of current optical correlation methods, is to assume a uniform displacement over a small window (typically between 3 and 100 pixels wide/high). However, this assumption can lead to wrong estimations, notably close to a sharp discontinuity such as a fault rupture. We present here the first data-based method to perform ground displacement estimation, relying on a machine learning model and a synthetically generated database. This database is used to train a model to retrieve the local displacement for a given image pair. It includes images containing synthetic sharp displacement boundaries in order to learn a more realistic machine learning model. Our results shows that we improve the accuracy near fault ruptures compared to state-of-the-art methods, which is important for studying the mechanics of near-fault processes.

**Key Points:**

- The first data-based optical correlation method, based on convolutional neural networks and generation of a synthetic training database
- It improves the displacement precision around sharp discontinuities (such as fault surface ruptures) compared to state-of-the-art methods
- It generalizes to different imagery (both Landsat-8 and Pléiades) and is computationally efficient (15 s for a 1024 × 1024 image)

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

T. Montagnon and J. Hollingsworth, tristan.montagnon@univ-grenoble-alpes.fr; james.hollingsworth@univ-grenoble-alpes.fr

**Citation:**

Montagnon, T., Giffard-Roisin, S., Dalla Mura, M., Marchandon, M., Pathier, E., & Hollingsworth, J. (2024). Sub-pixel displacement estimation with deep learning: Application to optical satellite images containing sharp displacements. *Journal of Geophysical Research: Machine Learning and Computation*, *1*, e2024JH000174. https://doi.org/10.1029/2024JH000174

Received 16 FEB 2024
Accepted 13 SEP 2024

**Author Contributions:**

**Conceptualization:** Tristan Montagnon, Sophie Giffard-Roisin, Mauro Dalla Mura, Erwan Pathier, James Hollingsworth
**Data curation:** Tristan Montagnon
**Formal analysis:** Tristan Montagnon



## 1. Introduction

### 1.1. General Context

Precise estimation of ground displacement at regional scales is fundamental for the study of natural hazards, such as earthquakes, volcanoes, landslides, as well as monitoring of glaciers (Heid & Kääb, 2012; Hollingsworth et al., 2012; Lacroix et al., 2019; Leprince, Ayoub, et al., 2007; Leprince, Barbot, et al., 2007; Van Puymbroeck et al., 2000). In the case of earthquakes, an accurate and unbiased estimation of ground deformation, especially in

**Funding acquisition:** Erwan Pathier, James Hollingsworth
**Investigation:** Tristan Montagnon
**Methodology:** Tristan Montagnon, Sophie Giffard-Roisin
**Project administration:** James Hollingsworth
**Resources:** Mathilde Marchandon
**Software:** Tristan Montagnon
**Supervision:** Sophie Giffard-Roisin, Erwan Pathier, James Hollingsworth
**Validation:** Tristan Montagnon
**Visualization:** Tristan Montagnon, James Hollingsworth
**Writing – original draft:** Tristan Montagnon, Sophie Giffard-Roisin, James Hollingsworth
**Writing – review & editing:** Tristan Montagnon, Sophie Giffard-Roisin, James Hollingsworth

the near-field of surface ruptures, is crucial to address the location, geometry, spatial distribution (on- vs. off-fault) of slip, and the down-dip slip distributions of the causative fault. Accurate characterization of the near-field displacement around surface ruptures, in turn, provides valuable constraints needed to understand the physics of earthquake slip (Marchandon et al., 2021), and to anticipate the seismic hazard posed to neighboring infrastructure and populations (Padilla & Oskin, 2023). Such observations are essential for addressing many areas of earthquake science, including fault mechanics, seismology, and structural/geological evolution of faulting, as well as earthquake engineering, including seismic hazard assessment, and seismic design. Recent observations have shown that the degree of slip localization may vary along fault ruptures in individual earthquakes, as well as between different earthquakes, and in different settings (Cheng & Barnhart, 2021; Dolan & Haravitch, 2014; C. Milliner, Dolan, et al., 2016). The mechanism by which slip localization varies in a surface rupturing earthquake isn't well understood, although several controlling parameters have been proposed (e.g., lithology, rupture velocity, earthquake magnitude, fault geometry, topography) (Ajorlou et al., 2021; Barnhart et al., 2020; Zinke et al., 2019). One difficulty in addressing this issue lies in the relative paucity of observational data currently available, rendering a comprehensive statistical analysis of the various parameters, and their trade-offs, challenging. Furthermore, of the modest number of case studies available, the observations are additionally complicated by the quality of the displacement maps obtained by image correlation. In particular, our ability to accurately resolve the degree of slip localization may be limited by the methodology used to generate the displacement map.

### 1.2. Optical Image Correlation

Optical satellite geodesy has revolutionized how we characterize ground deformation associated with natural hazards such as earthquakes (Avouac & Leprince, 2015; Marchandon et al., 2022). Optical Image correlation (OIC) is an imaging technique able to retrieve a full surface displacement field between two optical satellite images acquired over the same area at different times (separated by hours up to years). The main advantage of OIC, compared with more precise satellite geodetic techniques such as Interferometric Synthetic Aperture Radar (InSAR), lies in the ability to resolve large and high-strain displacements with very high spatial detail and using images with relatively long temporal baselines (up to many years). OIC has therefore been widely applied to the study of various sources of ground motion, such as surface rupturing earthquakes, volcanoes, landslides, and glacial movement (Leprince et al., 2008). In the study of earthquakes, where displacements are generally small relative to the pixel size, sub-pixel precision is critical for accurately capturing the displacement field, especially close to surface ruptures, where the deformation may become complex (e.g., sharp discontinuities, distributed off-fault deformation, secondary faulting, along-strike slip variability.). In this sense, earthquakes can represent a relatively challenging feature to measure with OIC, compared with large, fast-moving features such as glaciers.

In recent years, OIC has gained utility in constraining the near-field displacement around earthquake surface ruptures (Ajorlou et al., 2021; Antoine et al., 2021; Cheng & Barnhart, 2021; Michel & Avouac, 2002; C. W. Milliner et al., 2015; Scott et al., 2018; Zinke et al., 2019), in part due to the increasing availability of high resolution optical satellite data sets, as well as the difficulty in constraining these regions using InSAR (which typically decorrelates in areas with large strain gradients, e.g. due to cycle skipping and surface changes, (Lasserre et al., 2005; Pinel-Puysségur et al., 2022)). OIC can therefore provide valuable constraints on the localization (or distribution) of slip between the primary fault core and the neighboring (decimeter-scale) off-fault damage zone.

Over the past few decades, various OIC methods have been developed to quantify image displacements (Guizar-Sicairos et al., 2008; Hirschmuller, 2007; Nefian et al., 2009; Zach et al., 2007). Focusing on the approaches largely used in the remote sensing community (involving satellite and aerial imagery), these can be broadly grouped in correlation-based techniques working in the spatial (C. Milliner & Donnellan, 2020; Rosen et al., 2004; Rosu et al., 2015; Rupnik et al., 2017) or frequency (Fourier) domain (Leprince, Barbot, et al., 2007; Rosu et al., 2015; Tong et al., 2015; Van Puymbroeck et al., 2000). Spatial cross-correlation is a fundamental technique for estimating ground deformation by comparing reference and target images through a sliding window approach. This technique has been used (and refined) to efficiently solve the correlation problem (Rosu et al., 2015; Scambos et al., 1992). Frequency-based correlation is considered to be more accurate and faster (given that convolution in the spatial domain is equivalent to multiplication in the frequency domain) than commonly used spatial correlation methods such as normalized cross correlation (Tong et al., 2019). Both methods generally employ a block-based matching scheme, employing a sliding window to capture the local displacement field over the full image. However, this may lead to high frequency noise in the final displacement

map (at wavelengths similar to the sliding window dimension). Therefore, some approaches also make use of multi-scale regularization schemes to mitigate this effect (Leprince, Barbot, et al., 2007; Rosu et al., 2015).

Satellite images are subject to sources of noise leading to correlation bias or even temporal decorrelation. For example, the difference in illumination between two acquisitions, typically acquired with a difference of weeks to months, can introduce topographically correlated artifacts in the resulting estimated displacement map (Hollingsworth et al., 2017; Lacroix et al., 2019; Leprince, Ayoub, et al., 2007; Van Puymbroeck et al., 2000). Reflectance changes due to vegetation or anthropogenic changes between acquisitions may also hinder the correlation over long time periods (Van Puymbroeck et al., 2000). Noise present in the acquisitions can also be a source of errors in the correlation results. Low level image pre-processing carried out by the data providers, such as image resampling performed in the mosaicking of the raw acquisitions from single arrays for forming an image in staggered pushbroom sensors, or the projection in a geographic reference system can introduce additional errors such as periodic artifacts appearing at specific spatial frequencies (Leprince, Ayoub, et al., 2007; Tong et al., 2015). Spatial regularization schemes (Rosu et al., 2015), or frequency masking (Leprince, Ayoub, et al., 2007) have been used to help mitigate high frequency noise in the resulting displacement maps, although they can result in smoothing of the spatial detail. However, these approaches do not address limitations in the correlation process, such as the underlying assumption of rigid translations over the correlation window (Kanade & Okutomi, 1994), or spatial smoothing associated with the correlation kernel (C. W. Milliner et al., 2015), which may bias the displacements in the vicinity of sharp discontinuities typical with earthquake surface ruptures. In particular, smoothing of the displacement field in the vicinity of a sharp discontinuity will artificially distribute displacement into the neighboring areas. In the case of earthquakes, this effect may severely bias our estimates of slip localization, and the partitioning of slip between the fault plane and the surrounding damage zone, as well as estimates of near-field strain (Ajorlou et al., 2021).

### 1.3. Contributions

The main objective of this study is to create a neural network-based displacement estimation method, which to the best of our knowledge is the first attempt for the estimation of sub-pixel displacements in the context of ground deformation associated with natural hazards. We specifically aim to tackle the challenge of estimating sub-pixel displacements in the presence of discontinuities, in particular near fault ruptures, where limitations linked with the assumption of a rigid translation (Kanade & Okutomi, 1994) or the spatial smoothing (C. W. Milliner et al., 2015) are not explicitly addressed with literature.

We developed a pipeline using a convolutional neural network (CNN) to solve the sub-pixel displacement estimation problem. CNNs have the capability (a) to be robust to variations (noise and lightning for example); (b) to learn relevant features from the data without manual feature engineering, which simplifies the modeling process and enhances the ability to capture complex patterns; (c) to generalize well to unseen data with similar characteristics to the training set, making them suitable for tasks where correlation properties may vary. A novel aspect of the proposed technique is that we address the specific challenge faced when discontinuities are present within the sliding window, that is, in the near-field of fault ruptures. We achieve this by training a sub-pixel model on a realistic synthetic data set that includes samples mimicking real fault discontinuities (see Figure 1). Our implementation includes a two-step pipeline that can effectively estimate (a) significant displacements exceeding one pixel by initially employing a coarse model in a pixel-scale registration step, and subsequently (b) refining the sub-pixel details centered around the pixel-scale displacement using a sub-pixel model. Our technique is able to perform efficiently the estimation on a $1024 \times 1024$ pixels satellite images pair in 15 s, competing with state-of-the-art methods computation time (details in Section 5).

This paper contributes to the literature in three major aspects:

(1) The creation of large realistic synthetic training sets that allow data-based techniques to learn how to retrieve surface deformation with sub-pixel accuracy and precision.
(2) The development of the first end-to-end neural network-based OIC method, as a 2-step procedure: first a coarse estimation (larger than one pixel), then a finer estimation (sub-pixel refinement).
(3) Comparisons with state-of-the-art correlation methods COSI-Corr (Leprince, Ayoub, et al., 2007) and MicMac (Rosu et al., 2015) to quantitatively (with synthetic realistic images) and qualitatively evaluate the results. The latter case is addressed through a study of the 2019 Ridgecrest earthquake (Antoine et al., 2021; Ross et al., 2019), where we characterize the co-seismic displacement field using very high resolution Pleiades satellite data.

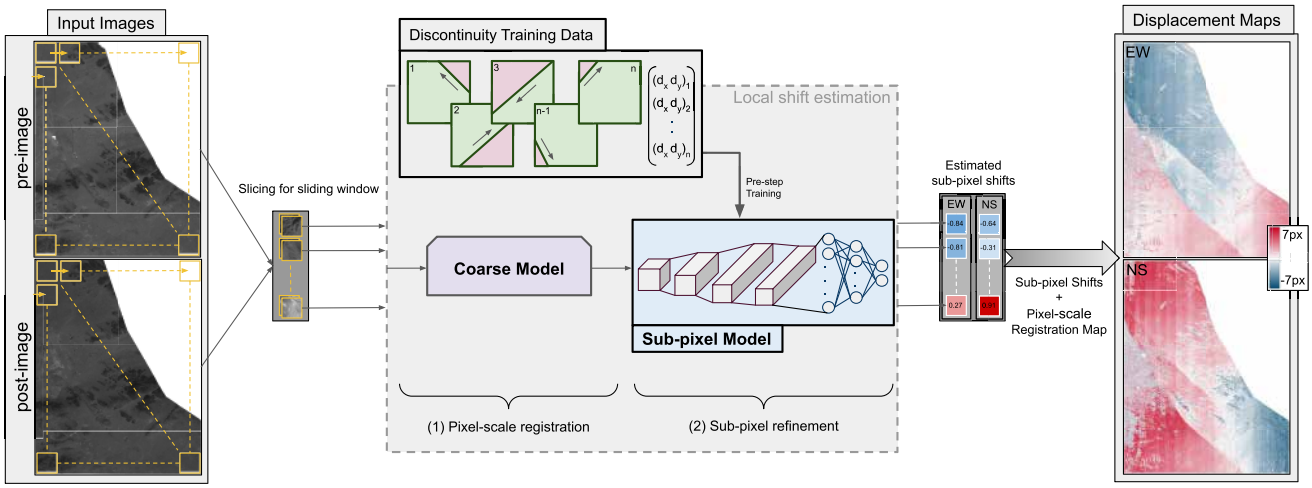


**Figure 1.** From left to right: A pair of input images (pre-event and post-event images) are acquired on two different dates. The sub-pixel displacement field between the two images is retrieved using first a coarse model registering every pair of windows pixel-wise, followed by the sub-pixel model that extract the sub-pixel shift; displacements is given in the row (North-South) and column (East-West) direction. Our sub-pixel model is trained with data containing discontinuities.

## 2. Related Works

OIC is based on image registration principles, in which the goal is to estimate a displacement (or disparity) between common features (represented by pixels) present within two images (Sutton et al., 2009). These two images are denoted pre-image (a reference image) and post-image (a second image, acquired after the reference), and we will use this terminology in this study. Sub-pixel OIC refers to the estimation of the displacement field between two images with a precision less than the image resolution (i.e., <1 pixel) (Pan et al., 2009). Attaining sub-pixel precision requires interpolating into the sub-pixel domain. In order to solve this problem, one common approach (in block-based matching schemes) is to assume a homogeneous displacement between the two images (i.e., the displacement is smooth enough to consider a locally rigid deformation). In most remote sensing cases, the transformation between the pre- and post-images is dominated by a simple rigid translation in 2-dimensions, and rotations or scale changes, etc., are assumed to be small. In these cases, we can thus simplify our registration problem to just 2 degrees of freedom.

### 2.1. Spatial Correlation Methods

A correlation measurement between two images of scalar values $I_1$ (reference) and $I_2$ (template) of same size consists of retrieving a similarity score, that indicates how well the content of the two images match and is typically used in template matching problems (Lewis, 1995). The cross-correlation $C_{I_1,I_2}$ between image $I_1$ and $I_2$ is computed by integrating within the image domain the pixel-wise product between the two images for each 2D spatial shift, represented by a translation vector of components $(x,y)$:

$$C_{I_1,I_2}(x,y) = \sum_n \sum_m I_1(x+n, y+m) \cdot I_2(n,m). \tag{1}$$

The estimation of the integer 2D shift between the two images can be performed by finding the couple $(x,y)$, denoted by $(x^*,y^*)$, that maximises the correlation score:

$$(x^*,y^*) = \underset{x,y}{argmax}\ C(x,y). \tag{2}$$

We recall that $(x^*,y^*)$ might not be unique and corresponds to a global translation of the two images, meaning that all pixels in the template image will get shifted by the same amount.

Traditional methods (Leprince, Ayoub, et al., 2007; Rosu et al., 2015) use the principle of cross-correlation to retrieve displacement maps that describe the surface displacement field between two satellite images in two directions: East-West (horizontal) and North-South (vertical). Let us assume we have two images $I_1$ and $I_2$, and

we want to retrieve the 2D displacement field between them. To produce a map of the displacement field based on cross-correlation, we consider a spatially local approach consisting of splitting the template image $I_2$ into a set of rectangular windows of a given size. Each window $\{W_{k,l}\}$ corresponds to the set of pixels of the image $I_2$ included by the window when centered at the $I_2$ coordinates $(k,l)$. We first compute the correlation coefficient map (function of $(x,y)$) for every $W$ with origin (e.g., the window center) at coordinates $(k,l)$, named $C_{I_1,W}(k,l,x,y)$:

$$C_{I_1,W}(k,l,x,y) = \sum_n \sum_m I_1(x+n, y+m) \cdot W_{k,l}(n,m). \tag{3}$$

Note that in practise, we do not compute the correlation coefficient map on the whole reference image $I_1$, but restrict the $(x,y)$ to a search area around the coordinates of $W$. Then, from $C(k,l,x,y)$, that is a 4-way tensor, we can compute $(x^*_{k,l}, x^*_{k,l})$, which correspond to the displacement vector maximizing the correlation score for every $(k,l)$:

$$\left(x^*_{k,l}, y^*_{k,l}\right) = \underset{x,y}{argmax}\ C_{I_1,W}(k,l,x,y). \tag{4}$$

From this, we are able to retrieve the displacement field $\Delta(k,l)$:

$$\Delta(k,l) = \left(k - x^*_{k,l}, l - y^*_{k,l}\right). \tag{5}$$

In the state-of-the-art sub-pixel spatial correlation approach of MicMac (Rosu et al., 2015), the displacement estimation is based on normalized cross correlation, in which the displacement precision is refined iteratively using a progressively smaller search space with each iteration. MicMac can be less sensitive to image noise, and to large spatial heterogeneities due to large time differences between the acquisitions (Rosu et al., 2015) than frequency methods. However, spatial correlation can be computationally expensive, as it involves a convolution process involving many computations. Working with high resolution data containing large displacements also necessitates the use of larger search spaces, further increasing the computation time. The correlation can also fail, or be heavily biased under particular noise conditions (e.g., changes in atmospheric conditions, sensor artifacts, radiometric differences, geometric distortions, stereoscopic artifacts, aliasing, color saturation) or very close to the fault, by the nature of the assumptions made.

### 2.2. Frequency-Based Correlation Methods

Another way to retrieve displacement between two images is to work in the frequency domain. The general principle of frequency-based correlation is to compute the normalized cross-power spectrum $Q_{I_1,I_2}$ (i.e., the complex conjugate of the Fourier transform $\mathcal{F}\{\cdot\}$ of an image multiplied element-wise by the Fourier transform of a second image), and retrieve its inverse Fourier transform $\mathcal{F}^{-1}\{\cdot\}$, that will give us the normalized cross-correlation matrix $R_{I_1,I_2}(x,y)$ for a specific template image $I_2$ compared with $I_1$:

$$R_{I_1,I_2}(x,y) = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}\{I_1\}\mathcal{F}^*\{I_2\}}{|\mathcal{F}\{I_1\}\mathcal{F}\{I_2\}|}\right\}. \tag{6}$$

The relative displacement between the two images is then determined spatially from the position of the peak in this correlation matrix (Guizar-Sicairos et al., 2008). To estimate a displacement field over the full image space, we split $I_1$ and $I_2$ into a set of template windows $W_1$ and $W_2$ of coordinates $(k,l)$ which are used to obtain the local displacement values $\Delta(k,l)$ from the local correlation coefficient maps $R_{W_1,W_2}(k,l,x,y)$.

$$R_{W_1,W_2}(k,l,x,y) = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}\{W_{1_{k,l}}\}\mathcal{F}^*\{W_{2_{k,l}}\}}{|\mathcal{F}\{W_{1_{k,l}}\}\mathcal{F}\{W_{2_{k,l}}\}|}\right\}. \tag{7}$$

Alternatively, for every set of windows, $\Delta(k,l)$ may be estimated directly in the frequency domain (Leprince, Ayoub, et al., 2007; Tong et al., 2015; Van Puymbroeck et al., 2000) from the 2-D slope of the normalized cross-spectrum's unwrapped phase.

In the sate-of-the-art frequency-based correlator, COSI-Corr (Leprince, Ayoub, et al., 2007), the displacement is estimated purely in the frequency domain. Iterative adaptive frequency masking helps to mitigate the impact of noisy (high) frequencies, while maintaining a reasonable computational cost (Leprince, Barbot, et al., 2007). As with all frequency-based methods, the technique gives increasingly robust results with larger correlation windows (Leprince, Ayoub, et al., 2007; Rosu et al., 2015), albeit at the expense of spatial detail. Therefore, correlation windows of ~32-by-32 (equivalent to an effective width of 16-by-16, accounting for the windowing function applied to reduce spectral leakage when computing the FFT) often represent the optimum trade-off between accuracy, acceptable noise levels, and spatial detail. Nevertheless, the assumption of a homogenous translation over this correlation window results in bias around sharp discontinuities or regions where the true displacement is more complex.

### 2.3. Data-Driven Image Registration

Convolutional neural networks (CNNs) (Krizhevsky et al., 2012; LeCun et al., 1989) have emerged as a powerful tool in all image processing fields. A CNN is able to *learn* a task from a set of training images in order to then *estimate* a prediction for new test images. The architecture of a CNN consists of a succession of different layers. First, the input layer takes as input one or more images. Convolutional layers apply a set of *learnable* convolution filters (known as kernels) to the input. Each filter scans over the input and computes a weighted sum of pixel values within its receptive field. Then, an activation function (commonly Rectified Linear Unit, or ReLU) is applied element-wise to introduce non-linearity into the network. Optional pooling layers reduce the spatial dimensions of the feature maps produced by the convolutional layers. Max-pooling layers retain the maximum value within a sliding window, effectively downsampling the outputs of convolutional layers (feature maps). The output layer produces the final estimations by linearly combining the outputs of the last layer. Training a CNN for a specific task such as image registration consists in updating the network's weights (convolution filters) iteratively with a specific gradient descent optimization algorithm to minimize the loss function, here measuring the dissimilarity between the network's predictions and the ground truth transformations. Therefore, training requires *labeled* data, which consists here of source and target image pairs (input) with known transformations (output, i.e. label). The data is typically split in 3 sets: training set for model learning, validation set for monitoring the performance during the training phase, and test set to assess the model's performance on unseen examples once the model is trained. We call *hyper-parameters* the set of variables controlling the deep learning model architecture and algorithm, as opposed to the *parameters*, or weights, which are optimized during training. The hyper-parameters are the number and size of layers, the optimization algorithm (e.g., Adam, SGD), the learning rate (step size at each gradient descent iteration), etc.

Image registration and displacement field estimation from optical images has been successfully addressed by recent data-driven approaches, and in particular CNNs, for example, in medical-imaging (Simonovsky et al., 2016; X. Yang et al., 2017), and remote sensing (Wang et al., 2018). In the domain of satellite images registration, CNNs have been shown to improve global image registration performance (Wang et al., 2018; Ye et al., 2018). Displacement field estimation between two images can also be efficiently solved by deep learning, for example, treating optical flow estimation as a learning task (Dosovitskiy et al., 2015). However, the large majority of registration problems focus on the estimation of large displacements (>1 pixel) from temporally dense data sets, while the estimation of sub-pixel shifts from temporally limited and distant acquisitions has been little studied. Several recent studies have demonstrated the potential of data-driven approaches to retrieve sub-pixel displacements (Boukhtache et al., 2021; Ilg et al., 2017; Montagnon et al., 2022; R. Yang et al., 2022). However, to the best of our knowledge, no application in remote sensing has yet been proposed for resolving sub-pixel displacements, particularly in the presence of local complexities such as sharp discontinuities, from optical satellite data, which are fundamental features for automatic slip characterization techniques.

## 3. Methodology

### 3.1. Approach and Pipeline of the Proposed Framework

Our proposed approach relies on the same principle as state-of-the-art OIC approaches. First, we work at the local scale with two small windows, $W_1$ and $W_2$, of size $k \times k$ (with $k$ the size of the sliding window in pixels; we use $k = 16$, as it has been demonstrated to be an effective window size from previous OIC studies of earthquake

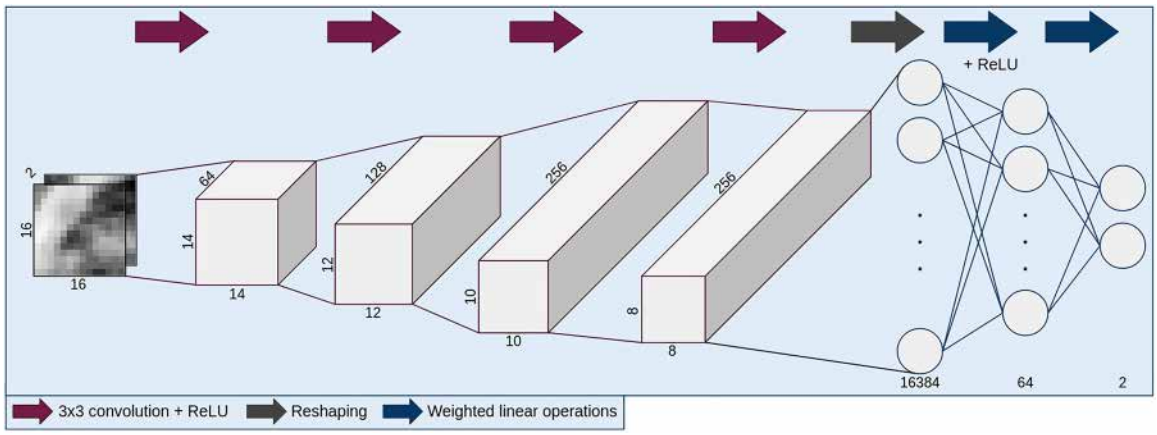


**Figure 2.** Architecture of the model used to achieve the sub-pixel estimation (cnn4l). The model takes the pre- and post-windows as input, and outputs the estimated shift (EW and NS displacements between the two windows).

deformation, and allows us to compare with state-of-the-art approaches; $k$ could be adapted). We also make the assumption of a locally rigid and non-rotating transformation, by evaluating the translational displacements $\Delta$ between the two windows. We develop an integrated 2-step pipeline able to estimate (a) the large scale displacements (>1 pixel) with a coarse model in a pixel-scale co-registration step, followed by (b) sub-pixel refinement centered on the pixel-scale displacement with a sub-pixel model. Both steps are implemented using a sliding window procedure. Figure 1 summarizes the full pipeline, taking two satellite images (global scale) as input.

### 3.2. Coarse Model and Sub-Pixel Model Architecture

Our approach consists of a two-step process to estimate shifts between image patches: a coarse estimation followed by a sub-pixel estimation.

First, estimating a shift at a pixel-level requires a larger patch window. For this initial step, we developed a model called cnn3l, which performs the first-step coarse co-registration of the sliding windows. The cnn3l model consists of 3 convolutional layers with 16, 32, and 64 kernels, respectively, each of size $5 \times 5$, and takes two $32 \times 32$ pixel patches, $W_1$ and $W_2$, as input.

After the coarse estimation, we refine the shift estimation to a sub-pixel level. For this purpose, we developed a 4 convolutional layer CNN architecture called cnn4l, which takes two $16 \times 16$ pixel patches, $W_1$ and $W_2$, as input. The model outputs a vector of two displacement values, representing the estimated shift (in fraction of pixels) between the two input windows, in the row and column directions (corresponding to north-south and east-west directions for UTM-projected imagery). The architecture of our network can be summarized as follows: the input is passed through four convolutional layers, with an increasing number of small kernels (64, 128, 256, and 256). The size of the kernels ($3 \times 3$) was selected to extract small features in already small ($16 \times 16$) windows and is well-suited to work on small displacements (Ilg et al., 2017). The output of each convolution is processed by the Rectified Linear Unit (ReLU) activation function. After the convolutions, two fully connected layers reduce the size of the data (from 16,384 to 64 to 2) and output the estimated shift. The structure and its parameters are summarized in Figure 2. Pooling operations are not used here; as the input images are already very small, the size reduction caused by the convolution operations without padding is sufficient to extract multi-scale features (the last feature maps shape is $8 \times 8$, thus, two times smaller than the input). The proposed architecture is generic and could be used with other window sizes than $k = 16$.

### 3.3. Generation of the Training Database

In the Earth Science community, no suitable archive of synthetic earthquake displacements currently exists for the purposes of training a Convolutional Neural Network. Furthermore, the lack of precise, spatially dense ground displacement measurements in real earthquake cases hinders the creation of a relevant database of suitable ground

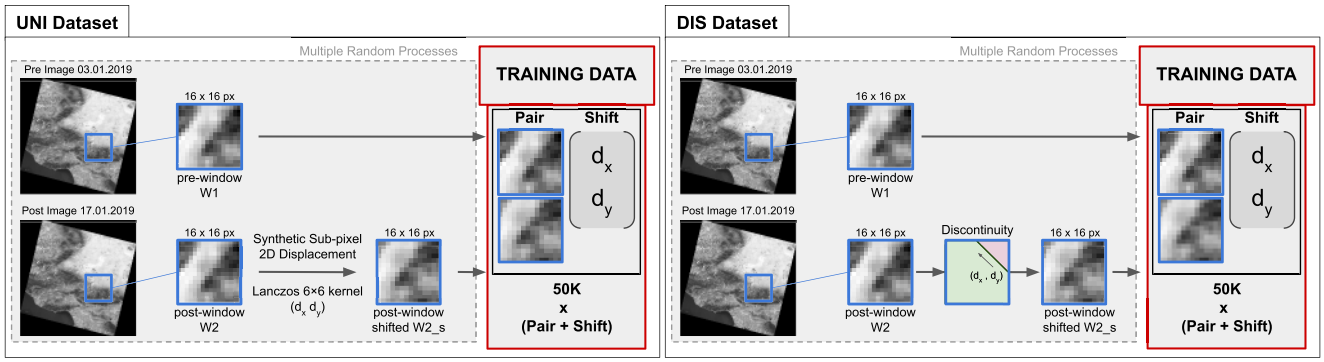


**Figure 3.** Creation of our two training sets, UNI (left) and DIS (right). For the UNI data set, the displacement incorporated in $W_2$ is uniform, when for the DIS data set, the deformation includes a discontinuity.

truth data. Consequently, we generate our own synthetic satellite imagery with known displacements, to train and validate our network. This training data set contains pairs of patches (input windows), together with the rigid deformation values linking them (targets). In order to create samples containing realistic reflectance noise (resulting from variations in illumination, changes in vegetation, etc.), we use real Landsat-8 acquisitions acquired over the same region, but on different dates. We select satellite images from areas which are stable between the two acquisitions (i.e., no ground displacement is present). Additionally, we globally co-register the two acquisitions using phase correlation (a Python implementation of the matrix-multiply DFT method of Guizar-Sicairos et al. (2008) available in the Scikit-image library), to reduce global mis-registrations that might remain after processing by the USGS.

We extract two windows, $W_1$ and $W_2$, respectively, from two large satellite images, $I_1$ and $I_2$, which are acquired over the same area on two different dates, $t_1$ and $t_2$. We build a synthetic displacement field, $\Delta$, used to warp $W_2$ to obtain $W_{2s}$, the distorted version of $W_2$. One sample of the training data set is the standardized pair $(W_1, W_{2s})$ (re-scaled with a 0-mean and a unit variance), with the target deformation $\Delta$ associated. A re-sampling algorithm is necessary during the warping process, because the shift applied is sub-pixel: we use Lanczos interpolation, with a kernel size of $6 \times 6$.

This can be summarized as follows:

$$W_{2s}(x,y) = f_{\Delta}(e_{\delta t}(W_1(x,y)))(x,y) \tag{8}$$

with $e$ the natural evolution of the ground during $\delta t$, and $f$ the distortion operation associated to $\Delta$. This process is repeated many times (with random $\Delta$, random window extraction location, random pair $(I_1, I_2)$, following a uniform distribution), to create a large number of unique samples. The procedure is summarized on Figure 3.

### 3.3.1. UNI Data Set

The first database that was generated assumes a perfectly uniform shift between the two patches, similar to the traditional correlation implementations. This uniform training data set is called "UNI." Let's fix: $\Delta = (\Delta_x, \Delta_y)$, where $-1 < \Delta_x, \Delta_y < 1$. Here, $\Delta$ simulates a uniform sub-pixel shift in $W_{2s}$. The UNI Data set is created with this uniform distribution of displacement fields, and is made of 125,000 samples for training (80% for training; 20% for validation) and 25,000 for testing.

### 3.3.2. DIS Data Set

A second training database was created in order to address the incorrect assumption of a uniform shift in cases where a displacement discontinuity is present within the sliding window. Here, $\Delta$ used to warp $W_2$ is not uniform, but contains a discontinuity (see Figure 3). This discontinuity data set is called "DIS." Formally,

**Table 1**
*Architectures Details (Number of Layers and Number of Feature Maps for Each Layer) and Mean Absolute Error (MAE) of the UNI Test Set (in Pixels, Mean Over Four Runs) for the Different Models Using* $16 \times 16$ *Input Patches*

| | cnn-1l | | | cnn-2l | | | cnn-3l | | | cnn-4l | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model name: | *s* | *m* | *l* | *s* | *m* | *l* | *s* | *m* | *l* | *s* | *m* | *l* |
| *conv1 (filters)* | 16 | 48 | 128 | 16 | 24 | 64 | 16 | 24 | 64 | 16 | 24 | 64 |
| *conv2 (filters)* | x | x | x | 20 | 48 | 128 | 20 | 48 | 128 | 20 | 48 | 128 |
| *conv3 (filters)* | x | x | x | x | x | x | 24 | 72 | 256 | 24 | 72 | 96 |
| *conv4 (filters)* | x | x | x | x | x | x | x | x | x | 28 | 256 | 256 |
| MAE (in pixels) | 0.127 | 0.137 | 0.133 | 0.107 | 0.106 | 0.106 | 0.110 | 0.110 | 0.106 | 0.114 | 0.111 | **0.105** |

*Note*. Notation: s: small; m: medium; l: large. The best number for the MAE is highlighted in bold.

$$\Delta(x,y) = \begin{cases} \Delta_a(x,y) = (\Delta_{x_a}, \Delta_{y_a}) & \text{if } (x,y) \in A \\ \Delta_b(x,y) = (\Delta_{x_b}, \Delta_{y_b}) & \text{if } (x,y) \in B \end{cases} \tag{9}$$

where A and B are the two areas (green and pink on Figure 3) created by intersecting a random line with the window square, with $area(A) > 1.05 \times area(B)$, and $(x,y)$ are the coordinates of the pixels in $W_2$. This means that the center pixel is always contained in A. Therefore, the target deformation for this sample is $\Delta_a$. With this simulated fault discontinuity, the model needs to identify the displacement of the largest area A in a given pair of windows, while some smaller area B is allowed to move in a different direction. Again, 125,000 samples are created for training (80% for training; 20% for validation) and 25,000 for testing.

We separately trained the cnn4l model with both UNI and DIS data sets, respectively giving two models: cnn4l-uni16 and cnn4l-dis16, both made to do the sub-pixel estimation. Two other training data sets were tested: (a) using $e_{\delta t}(W_1) = Id(W_1)$, that is, $t = 0$ using exactly the same acquisition for $W_1$ and $W_2$ ($W_1 = W_2$), and (b) using $e_{\delta t}(W_1) = Id(W_1) + n$ with $n$ a uniform random noise to train the model (adding artificial noise to replace the natural noise). However, the accuracy reached on our realistic synthetic data set (explained in the Section 4) and on real data (Ridgecrest study case) were unsatisfactory.

### 3.4. Sensitivity Study

In order to select the best hyperparameters, we evaluated our models on the test set (25,000 Landsat-8 $16 \times 16$ patches, acquired on different dates and locations, to guarantee that there is no common data used in both training and evaluation stages).

#### 3.4.1. The Sub-Pixel Model Architecture

We evaluated the sensitivity of our model with respect to the two main parameters controlling the architecture: the number of convolutional layers, and the number of filters in each convolutional layer. We evaluated the mean absolute error when using 1, 2, 3, and 4 convolutional layers, and with three different levels of filters per layer (small, medium and large). Table 1 gives details on the architectures tested, and quantitative precision on the evaluations. We trained the different architectures with the UNI training set, and tested on the UNI test set four times for each architecture, so every evaluation is a mean of four runs.

We see that the main precision gain is obtained with a depth of at least 2 convolutional layers. The best precision is reached with the most complex architecture: 4 layers with the largest number of filters. With this architecture, the model is complex enough to be able to use all the information of the training samples. Accuracy gain with more layers (+4 layers) was insufficient to justify the complexity and computational cost associated with a heavier CNN structure. This architecture was developed using the UNI data set, although the same trends are observed for the DIS data set, reaching 0.172, 0.132, 0.131, and 0.130 of MAE respectively for cnn-1l, cnn-2l, cnn-3l, cnn-4l, all with large number of filters. We thus retain this architecture for the rest of our study. However, future studies may explore broader architectures.

**Table 2**
*Mean Absolute Error (MAE) on the UNI Test Sets (in Pixels) for the Different CNN Models Taking as Input Different Window Sizes*

| | cnn4l-uni8 | cnn4l-uni16 | cnn4l-uni32 | cnn4l-uni64 |
|---|---|---|---|---|
| MAE (in pixels) | 0,130 | 0,105 | 0,0960 | 0,0884 |

To have a first sense of the performance of our two models cnn4l-uni16 and cnn4l-dis16, we can evaluate their accuracy on our DIS test set. Evidently, cnn4l-uni16 with a MAE of 0.163 performs worse than cnn4l-dis16 with a MAE of 0.130, that trained on the training set of DIS. Note that the MAE reached is lower than the MAE of cnn4l-uni16 on the UNI test set, due to the higher complexity of the DIS data set.

#### 3.4.2. The Window Size

The largest the images, the more information the model has to estimate the displacement. Therefore, on our synthetics (that have a uniform displacement), the best precision is achieved for the largest window size (Table 2). However, for real cases where the displacement is heterogeneous within an image, using large windows will bias the estimation. It is thus preferable to use a window size as small as possible. All trained models perform well (Table 2); however, for size: 8 pixels, the accuracy drops by 30% compared to size: 16 pixels. We will focus the rest of our study on the $16 \times 16$ window size, known to be more accurate than larger resolutions on real (i.e., not uniform) deformations.

### 3.5. Training the Pixel-Scale Co-Registration Model

The above trained cnn4l model captures shifts that are smaller than 1 pixel in each direction. Even though Earth deformations are contained in this range for the majority of remote images acquired, larger shifts can happen. We thus trained our cnn3l model, called cnn3l-5px, with another data set UNI-5px, very similar to UNI, but with $-5 < \Delta_x, \Delta_y < 5$ and $k = 32$. With this model, we co-register pixel-wise the sliding windows as a pixel-scale co-registration step (with a maximum estimation of 5 pixels), and then apply our sub-pixel model to refine our estimation. While the results of cnn3l-5px are more than satisfactory for estimating the rounded pixel-level shift (MAE = 0.37 pixels, notably smaller than 0.5 pixel), there is a clear advantage to perform the refinement with the specific sub-pixel cnn4l model. This procedure is detailed on Figure 1, and allows an estimation of large and small displacements. If we want to process images containing larger displacements ($\gg$5 pixels) with our cnn3l model, another training data should be built, with $|(\Delta_x, \Delta_y)| > 5$. Another solution would be to use a faster external existing method to perform the first step, that is, the less challenging pixel-size registration, and then refine the estimation with our sub-pixel model. Our pipeline will incorporate this flexibility.

### 3.6. Implementation Details

Our cnn4l-uni16 and cnn4l-dis16 models were implemented using PyTorch library in Python. The cnn4l-dis16 has reached convergence in 200 epochs (see Figure S1 in Supporting Information S1), while for cnn4l-uni16 only 99 epochs were needed to reach convergence (see Figure S2 in Supporting Information S1). We used early stopping, with a patience (i.e., the number of epochs to wait before early stop if no progress on the validation set) of 12, with an initial learning rate of $1 \times 10^{-3}$, decaying every 10 epochs with a factor of 0.8. We used the Adam optimizer with a weight decay of $5 \times 10^{-4}$, and the Mean Squared Error (MSE) loss. We chose a batch size of 128, averaging 10 min of computing per epoch for both models. Computations were performed on a GPU NVIDIA Tesla V100 NVLink. The pipeline is implemented in Python (a combination of Numpy and PyTorch).

We also make use of custom transforms (horizontal and vertical flips and n $\times$ 90° rotations) during the training to augment the data set.

## 4. Experiments

### 4.1. Evaluation on Realistic Synthetic Earthquake Images

In the absence of substantial ground truth data from real earthquakes, we first generate realistic synthetic earthquake images using Landsat-8 satellite acquisitions re-sampled to include realistic synthetic displacement fields; we then quantitatively validate our model and compare with existing state-of-the-art correlators using these synthetic satellite images. This evaluation data set is distinctly different from our training data set. First, it features a larger scale, with images of $1024 \times 1024$ pixels instead of $16 \times 16$ pixels. Additionally, the synthetic displacements included in this data set are significantly more complex and realistic (see Section 4.1.1 for details).

#### 4.1.1. Generation of Realistic Synthetic Earthquake Images

We first develop a pipeline that randomly generates realistic fault discontinuities with rough geometries and slip distributions embedded in a homogeneous elastic half-space (assumed to behave in a linear elastic fashion). Using analytical expressions, relating slip on triangular fault patches (in the elastic material) to displacement at the surface, we then compute the resulting surface displacement field produced by the prescribed fault geometry and associated earthquake slip distribution (Meade, 2007; Nikkhoo & Walter, 2015). The faults obey a length-displacement scaling consistent with natural earthquake faults (Scholz, 1982), rupturing only the seismogenic part of the crust, with geometric roughness consistent with natural faults (self-affine scaling, with a Hurst exponent of ∼0.8, (Candela et al., 2012)), fractal slip distributions (Amey et al., 2017; C. Milliner, Sammis, et al., 2016), and the potential for reduced slip in the uppermost crust (shallow slip deficit, (Fialko et al., 2005)). The length of the fault can range from 5 to 150 km, with a downdip width of 5–17 km and a strike of 0–360°. The fault dip varies between 30 and 90°. The fault models are discretized with an unstructured meshing approach (Mesh2D, (Engwirda, 2014)) using triangular displacement elements (TDEs), with sub-pixel resolution at the surface, which increases to several kilometers at depth; this depth-dependent spatial sampling accounts for the rapid decay of static displacements with distance to the source $\left(1/r^2\right)$, which prevents fine spatial details at depth to be resolved by surface measurements alone (i.e., through inverse modeling—see Page et al., 2009), while also significantly reducing the number of TDEs when computing the surface displacement field (generated using cutde, a GPU-accelerated implementation of Nikkhoo and Walter (2015)). Nevertheless, the high resolution of the fault model near the surface allows us to produce realistic displacements at the resolution of the satellite imagery. Finally, we also generate observation points (locations where we calculate a surface displacement) using an unstructured mesh, densifying points with proximity to the surface rupture, thereby concentrating the number of observation points to where there is significant variation in the displacement, further reducing redundant calculations with cutde; the irregular grid of output displacements are subsequently resampled onto a regular grid using the griddata function within the Python library: SciPy.

We then extract $1024 \times 1024$ patches from 15 m resolution Landsat-8 satellite images acquired on two different dates, but for a stable region in which no ground deformation has occurred. We perform an initial global co-registration step (using phase correlation), to ensure there is no mis-registration between the two images. The displacement field $\Delta$ is then used to warp the second (post) satellite image using a quintic-order spline re-sampling algorithm (Wiemker, 1996), to simulate an earthquake where the displacement field is fully known. The precision of this resampling approach (∼1/100th to 1/100th pixel) is significantly higher than the precision of state-of-the-art correlators (<1/10th pixel). Three different $1024 \times 1024$ images were created for evaluation; the creation of the input images is summarized on Figure 4.

#### 4.1.2. Comparisons With MicMac and COSI-Corr

The pre- and post-images are then compared using our new deep learning approach, COSI-Corr, and MicMac, to retrieve the displacement map. We then compute the absolute residual maps (using the ground truth displacement). In these examples, all displacements are kept in the range [-1,1], so we just focus on sub-pixel performance.

Using a sliding $k \times k$ pixels window with a stride of 1 pixel to have the best spatial resolution as possible, we compute the full displacement maps for the three evaluation examples using our two sub-pixel models, cnn4l-uni16 and cnn4l-dis16, as well as MicMac and COSI-Corr. For consistency, COSI-Corr was applied with a $32 \times 32$ window, because the effective correlation window width is reduced ∼half due to the windowing function applied (Heid & Kääb, 2012). MicMac typically uses smaller window sizes (the default is $9 \times 9$). However, we increase this to $15 \times 15$ to allow a more meaningful comparison with COSI-Corr and our pipeline (Although COSI-Corr can use $16 \times 16$ windows, thus yielding an effective window of $8 \times 8$, the displacements start to be biased too low). We also remove the spatial regularization option in MicMac, to keep consistency with our model and COSI-Corr (however, noise is still mitigated in MicMac by using a non-linear cost function, and in COSI-Corr by using adaptive frequency masking). Thus, we focus on the core sub-pixel performance of the various algorithms.

On Figure 5, the four EW and NS displacement maps are presented for three different realistic earthquake examples. To first order, the four models have similar results, with some small differences in texture, and in the sharpness of the fault boundary. Stacked displacement profiles spanning the fault in the first earthquake example

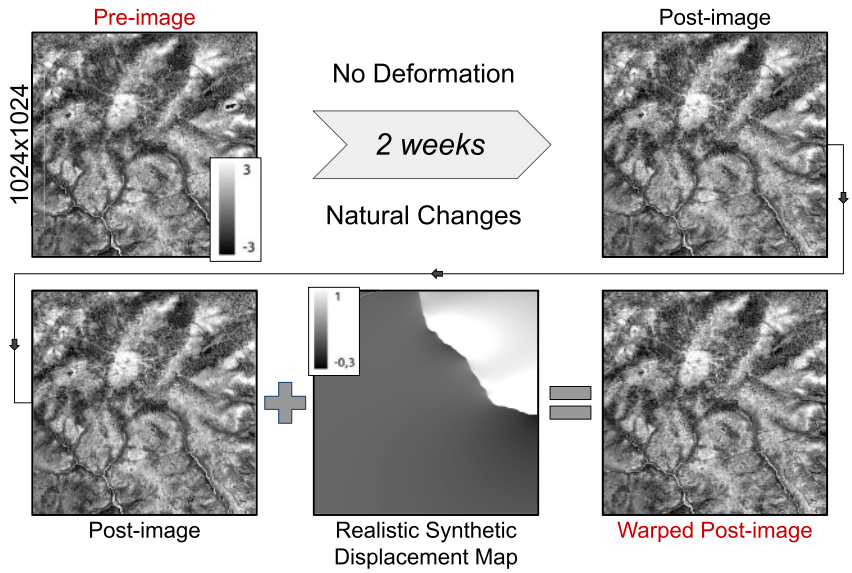


**Figure 4.** Creation of a pair of realistic synthetic earthquake images. Here, the warped post-image (bottom right) contains natural differences in reflectance due to the different acquisition time of pre-image (top left), and carries the synthetic displacement map (bottom center), that our model should retrieve. The values are in intensity (standardized) for the satellite images, and in pixels for the displacement map (E-W shown here).

(Figure 7, computed for the EW component of example 1) highlight a much sharper discontinuity (i.e., narrower displacement width) for our cnn4l-dis16 model, which compares most accurately to the ground truth of all the models (see Figure 7b); even compared to the default MicMac parameters ($9 \times 9$ with 0.3 regularization) or MicMac using the smallest available window size ($3 \times 3$ with 0.3 regularization, Figure 6; note that such parameters yield noisier correlation maps than we obtain with our CNN approach). MicMac without regularization (not represented here) is virtually identical to MicMac with regularization, because the stacked profiles suppress the high frequency noise, which would reflect the difference between the two profiles. This sharper, more accurate representation of the fault trace is also clear from the 2-D residual maps (Figure 8, computed for a portion of the EW component of example 1).

The mean residuals (MAE) for each of the 4 methods over the full extent ($1024 \times 1024$) of the displacement map, as well as a narrow zone extending 16 pixels from the fault trace are given in Table 3. The global mean errors of the four models are similar, ranging from $6.89e-2$ to $9.19e-2$ pixel. Our models perform slightly worse than MicMac and COSI-Corr (MAE increases by up to $2.30e-2$ pixel). This can be seen visually in the displacement maps (Figure 5): our cnn4l-uni16 and cnn4l-dis16 models yield displacement maps with slightly more high frequency noise. This is in part because we have made no effort to mitigate high frequency noise, unlike MicMac and COSI-Corr (simple post-processing such as Total Variation—L1 smoothing (weight 0.2)) (see Chambolle, 2004) allows our results to attain comparable level of smoothness with MicMac, while preserving sharper fault discontinuity (see Figure 6). Nevertheless, the stacked displacement maps spanning the fault displacement yield very similar displacements (Figure 7).

The error in the near-field of the fault trace is higher than the global residuals in each case. However, for cnn4l-dis16, the mean residual is reduced by 32% compared to the other methods (MAE drops by $\sim 8.1e-2$ compared to MicMac and $\sim 6.1e-2$ compared to COSI-Corr), reflecting a substantial improvement in resolving the displacement close to the fault rupture. The standard deviation and median around fault shows also a drop compared to MicMac and COSI-Corr. Also, we observe patches of $4 \times 4$ pixels around the discontinuity in MicMac results, due to the iterative refinement algorithm, that ultimately biases the precision close fault.

### 4.2. Evaluation on Real Case: Ridgecrest

On 4th and 7th July 2019, a large foreshock-mainshock earthquake sequence struck the town of Ridgecrest, in the Mojave desert region of southern California (Barnhart et al., 2019; Magen et al., 2020; Ross et al., 2019). The

10.1029/2024JH000174

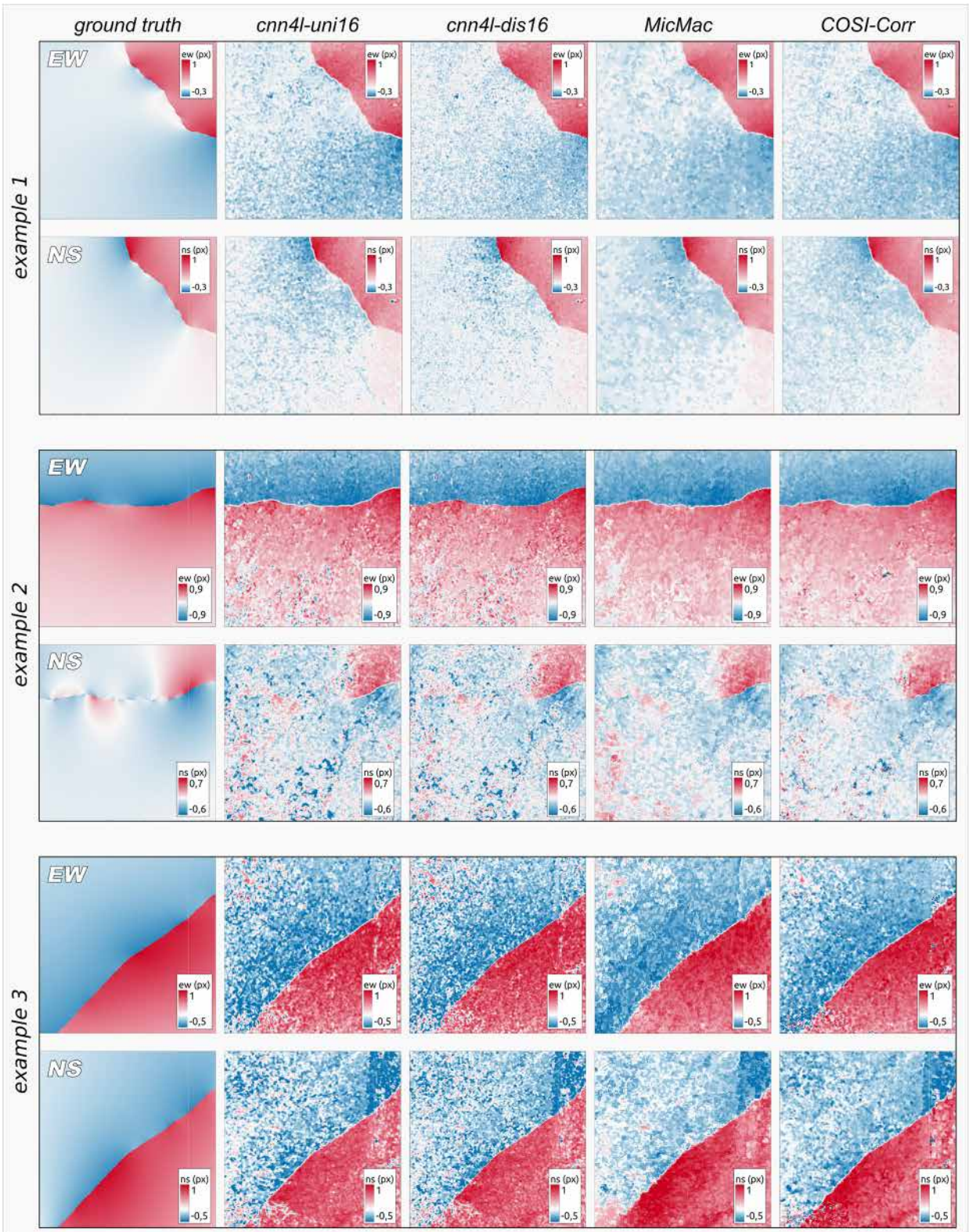


**Figure 5.** EW and NS displacement maps computed for 3 synthetic examples, for the four models cnn4l-uni16, cnn4l-dis16, MicMac, and COSI-Corr. On the left, the synthetic displacement maps used to warp the two satellite images. Results are expressed in pixels.

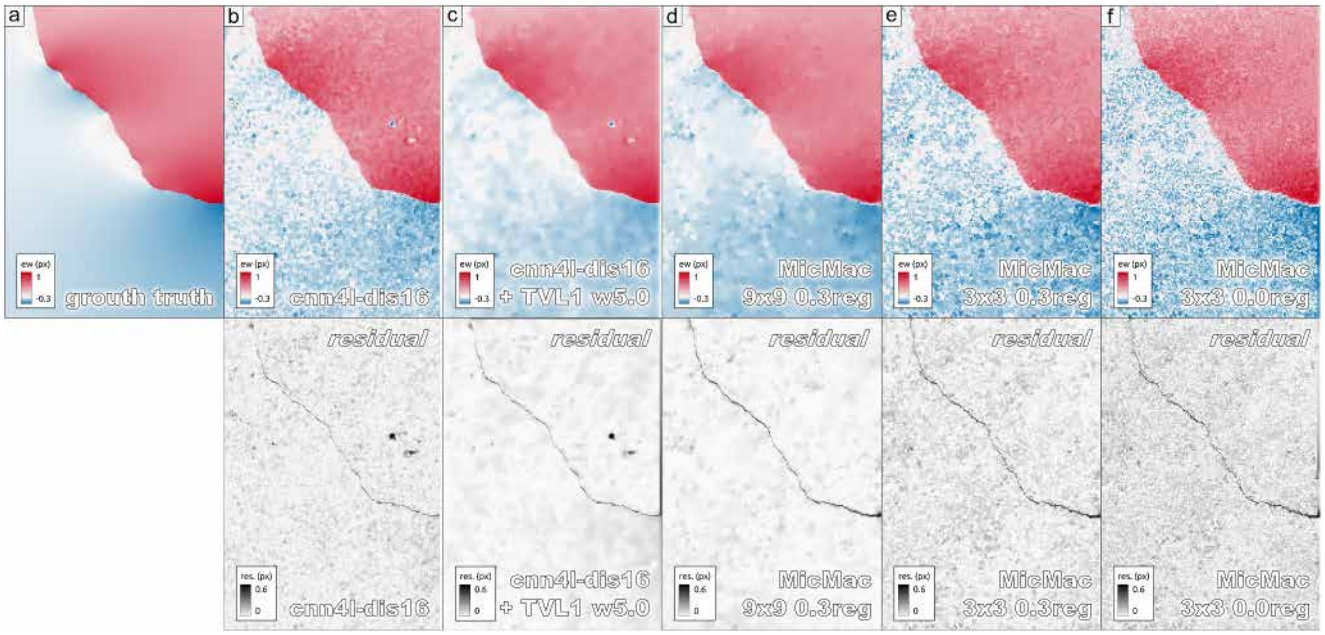


**Figure 6.** EW displacement maps and residuals for the (a) ground truth, (b) cnn4l-dis16, (c) cnn4l-dis16 with TVL1 smoothing (*skimage.restoration.denoise_tv_bregman*, weight 5.0), (d) MicMac (9 × 9 window, 0.3 reg), (e) MicMac (3 × 3 window, 0.3 reg), and (f) MicMac (3 × 3 window, 0 reg). Inclusion of simple TVL1 smoothing on our cnn4l-dis16 model highlights the potential of fast and simple post-processing approaches that may be used to reduce high frequency noise, giving smooth displacement maps comparable to default MicMac parameters, yet retaining a lower near-field bias. Lower panels show the absolute displacement residuals (relative to the ground truth).

foreshock broke a NE-SW-striking left-lateral fault, while the subsequent mainshock broke a conjugate NW-SE-striking right-lateral fault. Using very high resolution satellite imagery spanning both earthquakes, various studies used OIC to characterize the near-field displacements close to the surface rupture, and assess the extent of off-fault deformation associated with the event (Antoine et al., 2021; Barnhart et al., 2020; Magen et al., 2020; C. Milliner et al., 2021). These displacement maps also allowed an assessment of the local strain field, which turn helps to provide new constraints on the mechanics of fault slip in surface rupturing earthquake, and the distribution of slip between the fault core and neighboring damage zones.

Comparisons of our model with COSI-Corr and MicMac for optical data spanning the Ridgecrest earthquake is valuable for two reasons: (a) to test our model on satellite images containing real earthquake displacements, and (b) to test our model on different imagery to that used to train our model. In this case, our models are trained with Landsat-8 imagery (15 m), while we test on pre- and post-event Pleiades satellite images (from Antoine et al. (2021); resampled here from 0.5 to 1.5 m for memory constraints); thus, we assess the performance of our model on data acquired with a different sensor (with different characteristics, e.g. sensitive to different spatial wavelengths), different incidence angles (Pleiades imagery is typically acquired off-nadir, unlike Landsat-8), and at highly different resolutions (spanning ~1 order of magnitude). Furthermore, we avoid any artifacts potentially introduced by warping the satellite imagery, and we are not constrained by an artificial synthetic earthquake displacement model (however realistic we attempt to make it). Finally, the displacements involved during the earthquake can reach a magnitude of ~6 m, while the input imagery is 1.5 m resolution; therefore, we can test our full pipeline, which solves for displacements at the pixel-scale and then the sub-pixel scale (Figure 1). These comparisons therefore allow us to assess the generalizability of our technique. However, we note that for real cases of earthquakes, like Ridgecrest, we do not have a ground truth map of known displacements of similar density. Comparing different techniques is nevertheless useful to check consistency with existing and more established OIC methods. Furthermore, visual inspection of the displacement field close to the fault rupture can often reveal errors or bias which are identifiable from their spatial distribution, and which offer another means for comparison.

We compute displacement maps in 3 different areas using (a) our CNN-framework (step-1: pixel-scale registration with cnn3l-5px using 32 × 32 windows, followed by step-2: sub-pixel refinement with cnn4l-dis16 using

10.1029/2024JH000174

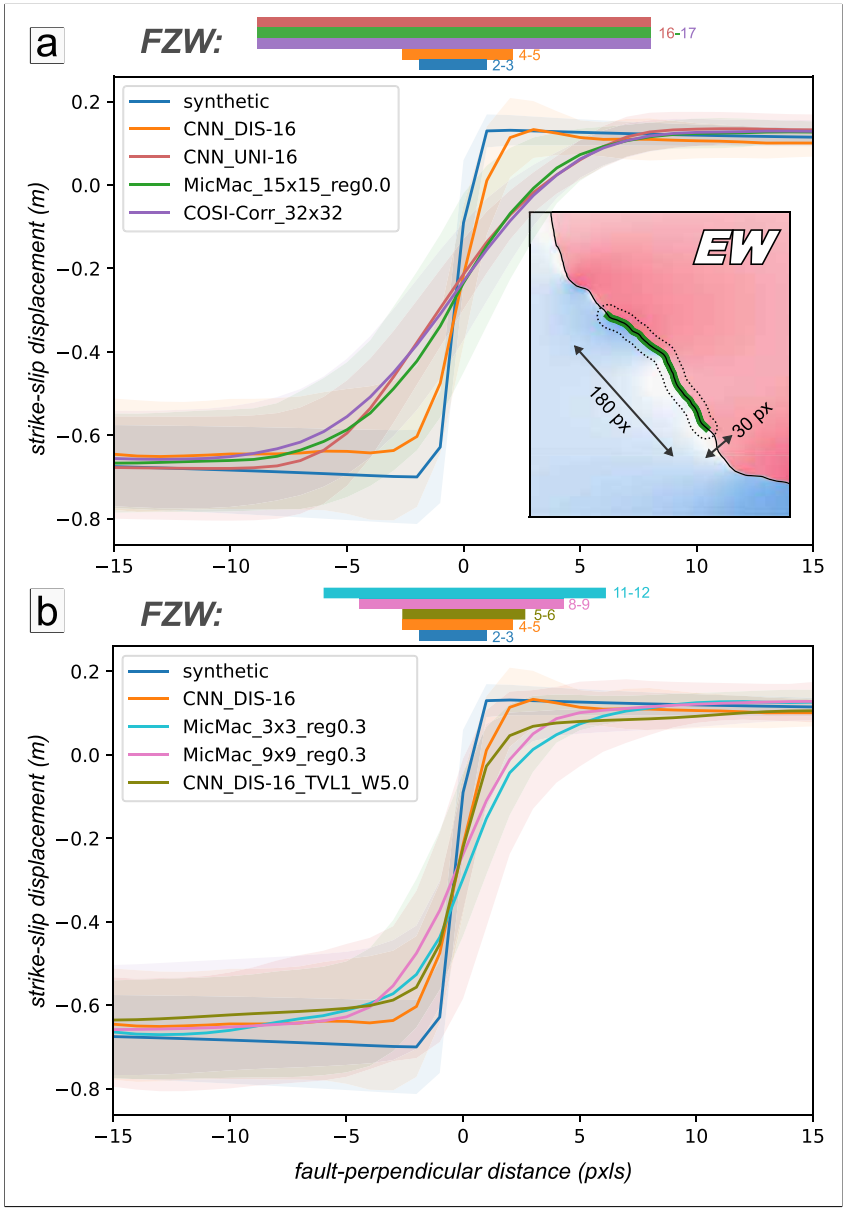


**Figure 7.** Stacked displacement profile on a portion of the fault for (a) the synthetic ground truth displacement field, the cnn4l-dis16 and cnn4l-uni16 models, COSI-Corr, and Mic-Mac (15 × 15, no regularization), and (b) MicMac (9 × 9 and 3 × 3 with 0.3 regularization, i.e. default parameters), and cnn4l-dis16 with TV-L1 smoothing (*skimage.restoration.denoise*$_t v_b$*regman*, weight 5.0). The fault zone width (FZW) is estimated in each case (colored bars, width given in), based on the approximate inflection points of the horizontal displacements each side of the fault. Note, a profile for MicMac with 0 reg (3 × 3 window) is virtually indistinguishable from the 0.3 reg case. The similarity results from the stacking, which reduces the impact of high frequency noise in 0 reg case; however, the higher level of noise is clearly visible in the 2D displacement maps (i.e., compare Figure 6e vs. Figure 6f). Inset in (a) shows the location of the stacked profile (stack width: 180 pixels; profile length: 30 pixels), relative to the EW displacement component (same as Figure 6a).

16 × 16 pixels windows), (b) COSI-Corr (32 × 32 windows), and MicMac (15 × 15 windows). A stride/step of 4 pixels is used for the CNN-pipeline and COSI-Corr at large scales (Figures 9a and 9c), and a stride of 1 for finer scales (Figures 9band 9d); MicMac (Figures 9e and 9f) can only stride at 1 pixel increments. The displacement maps for the first area are shown in Figure 9; the pre-image of this example was acquired on 23rd June 2012, and the post-image on 8th September 2019 (Antoine et al., 2021). Two additional results for other areas are shown in Figure S3 of the Supporting Information S1.

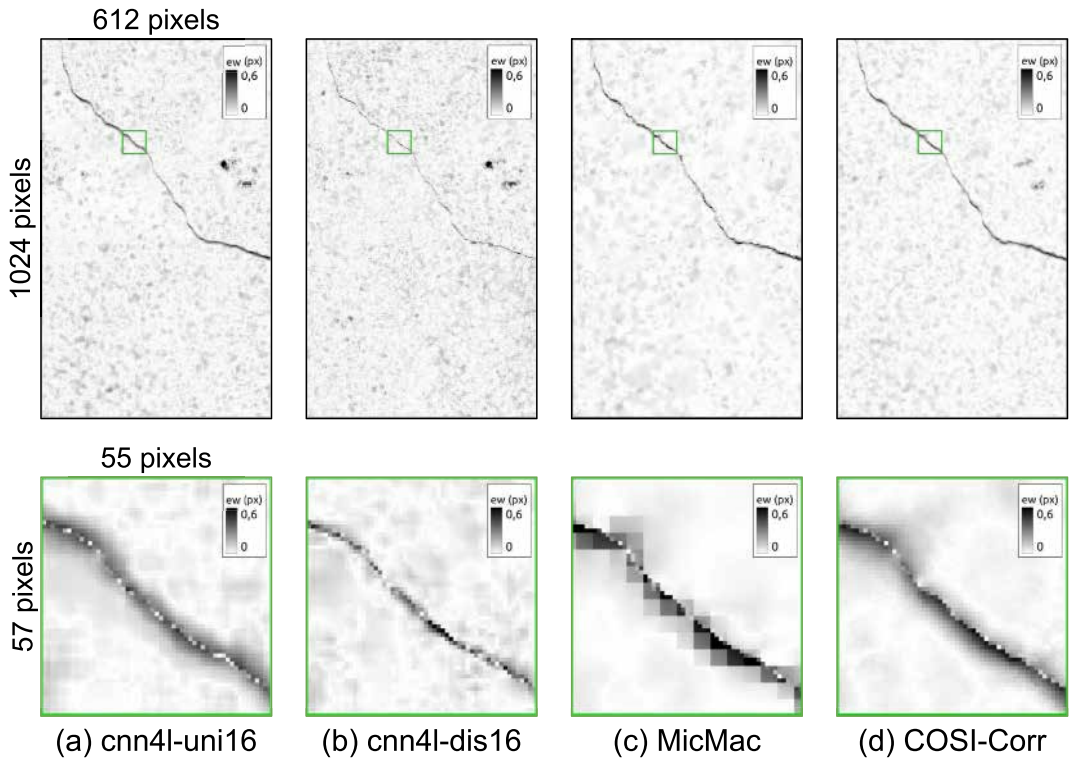


**Figure 8.** Residual maps of the displacements from Figure 5 (EW component of example 1) for cnn4l-uni16 model (uniform), cnn4l-dis16 (with discontinuities), MicMac and COSI-Corr.

All four methods give very similar displacement values globally, with small differences in the high frequency content (Figure 9). Both MicMac (and to a lesser extent COSI-Corr) give slightly smoother results, and in several small areas smooth over high frequency noise or regions of decorrelation (e.g., associated with buildings or roads), compared with our CNN pipeline. Nevertheless, globally the displacements are consistent with each other at medium and long wavelengths (Figure 9).

Close to the fault rupture (Figures 9b, 9d, 9f, and 10), where displacements are dominated by high-frequency components (i.e., sharp discontinuities), we see more significant differences between the various methods, consistent with our observations for the synthetic earthquake cases (e.g., Figure 8 bottom panel). MicMac yields a relatively clean, albeit slightly pixelated displacement map (Figure 9d), which is generally consistent with our CNN-pipeline result (Figure 9b). COSI-Corr yields a slightly smoother, lower resolution displacement map (Figure 9f), with occasional outliers, and apparent fattening/adhesion artifacts (Blanchet et al., 2011) along the rupture trace, indicating an artificially induced component of geometric roughness (which may result both from the frequency masking approach used to mitigate noisy frequencies, and the failure of a uniform translation to

**Table 3**
*Mean (Mean Absolute Error), Standard Deviations and Medians of cnn4l-uni16, cnn4l-dis16, Mic-Mac and COSI-Corr Residuals on the Three Synthetic Simulated* $1024 \times 1024$ *Maps Global and Around Fault*

| | cnn4l-uni16 | cnn4l-dis16 | MicMac | COSI-Corr |
|---|---|---|---|---|
| Mean residual | 0.0919 | 0.0909 | 0.0716 | **0.0689** |
| Standard Deviation | 0.0915 | 0.0875 | **0.0681** | 0.106 |
| Median | 0.0675 | 0.0678 | 0.0578 | **0.0487** |
| Mean residual around fault | 0.208 | **0.150** | 0.231 | 0.211 |
| Standard Deviation around fault | 0.181 | **0.153** | 0.215 | 0.237 |
| Median around fault | 0.157 | **0.104** | 0.163 | 0.140 |

*Note*. The best number for each category is highlighted in bold.

10.1029/2024JH000174

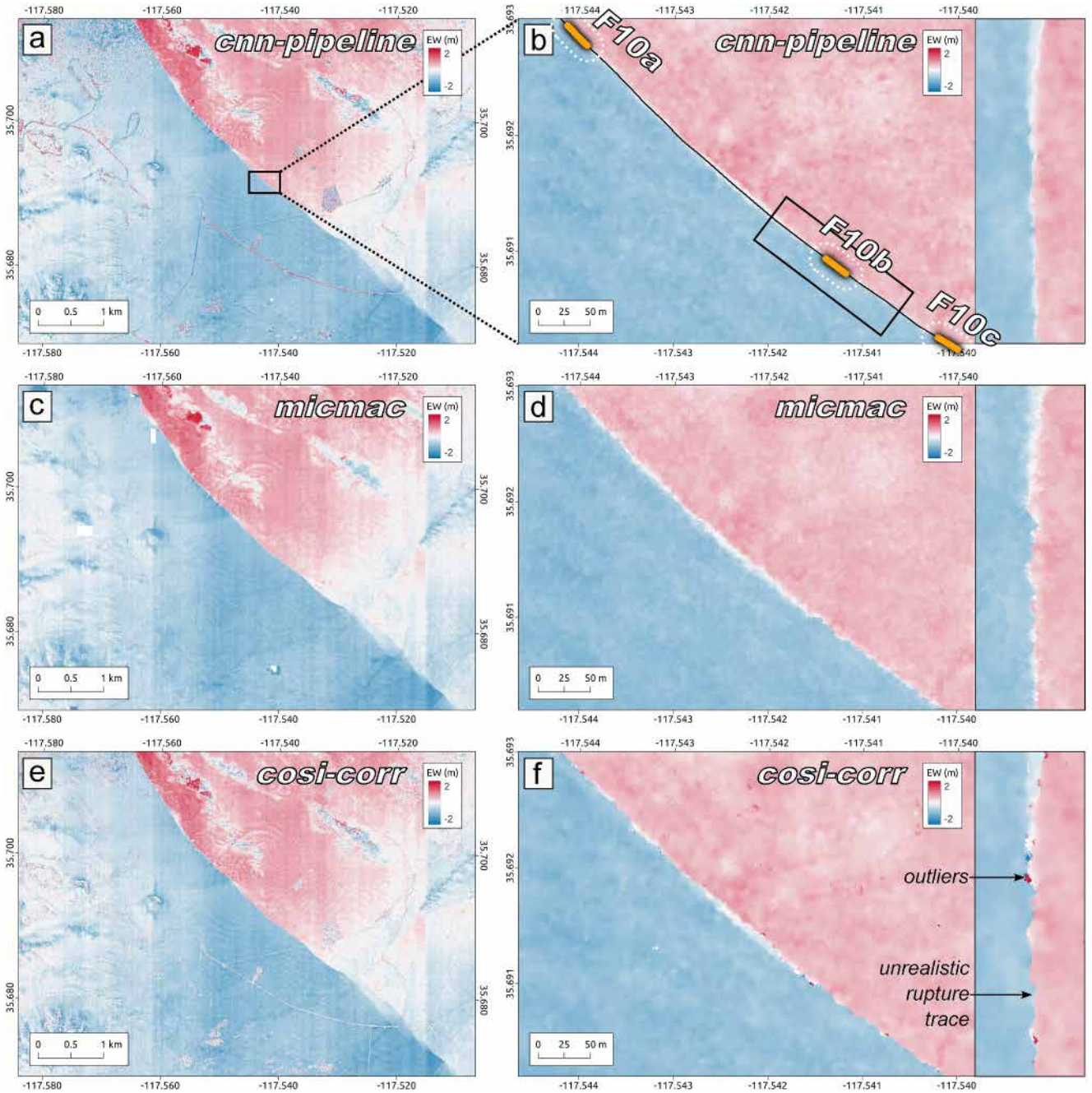


**Figure 9.** EW component of the displacement maps from Pleiades images for the Ridgecrest earthquake. (left-column) (a) Our 2-step CNN-pipeline, (c) MicMac, and (e) COSI-Corr. Positive pixel values indicates movement to the south and negative to the north. (right-column) Close-up region immediately around the fault rupture for (b) CNN-pipeline, (d) MicMac, and (f) COSI-Corr. Inset shows a further zoom of the rupture details. Orange lines in (b) show the stacked displacement profile locations in Figure 10.

describe a displacement field containing a discontinuity). Therefore, in real case examples, our CNN-pipeline seems to feature less bias in the near-field of the fault rupture.

Displacement profiles at three separate locations (Figure 10) all show a sharper displacement gradient more accurately located across the fault rupture (centered on 0 pixels along the x-axis) for the CNN-pipeline, highlighting the ability of our approach to capture sharp displacements and minimizing smoothing across the fault. MicMac (using both a 15 × 15 window with no regularization, and the default settings of 9 × 9 window 0.3

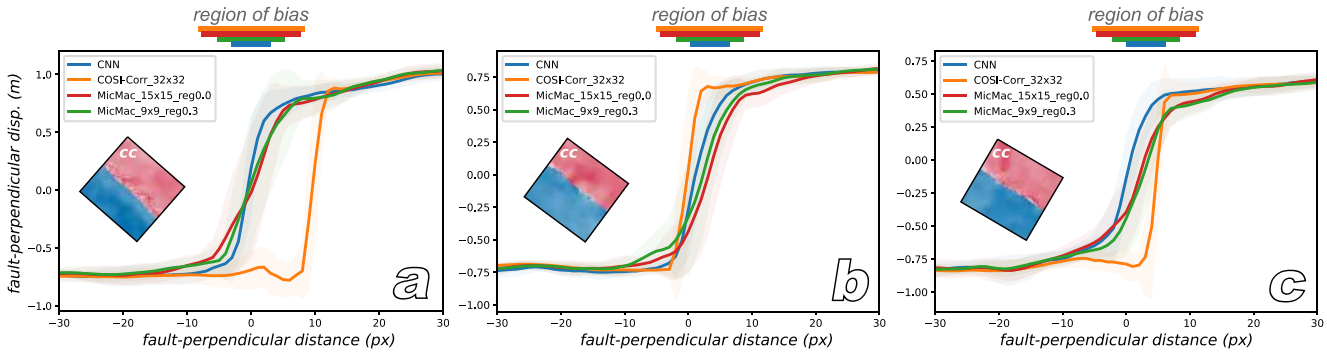


**Figure 10.** Fault-parallel displacement profiles across 3 sections (a–c) of the Ridgecrest rupture shown in Figure 9. In each case, the rupture is located at 0 pixels along the x-axis, and is constrained by field mapping by the USGS (Ponti et al., 2020). Each profile is a stack of 40 profiles. The approximate zone of near-field bias for each correlation approach, estimated from the correlation window size, is shown by the colored bars. Inset figure shows a zoom of the EW displacement data for the COSI-Corr approach used in the stack; high frequency noise close to the rupture is clearly visible in each case (compare with CNN-pipeline and MicMac, where this is not the case—see Figure 9).

regularization) gives similar results to our CNN-pipeline, albeit with slightly more smoothed discontinuities. Meanwhile, COSI-Corr gives apparently very sharp displacements, although the location is often far from those of MicMac and the CNN-pipeline, and not centered on the true fault location at 0 pixels, which is based on mapping by the USGS using field and remote sensing data (Ponti et al., 2020), and thus reflects the best independent ground truth we have regarding the location of the rupture. Therefore, the details of the COSI-Corr result in the near-field, especially the sharp displacement gradients, are not robust. This is also reflected visually in the wavy fault trace, and high frequency noise visible along the rupture in Figure 9 and inset figures of Figure 10. The distance over which COSI-Corr is inaccurate lies within ±8 pixels of the true rupture location (thus, a biased region of 16 pixels); this distance is consistent with the effective window size of 16 × 16 pixels for a 32 × 32 window (once the windowing function is applied). The larger correlation window of COSI-Corr thus increases the distance over which near-field displacements are corrupted.

The main interest of these figures are to demonstrate that we are very similar to state-of-the-art methods. Future efforts will re-examine the near-field displacements for the Ridgecrest earthquake using more comprehensive analysis combining existing satellite imagery with our CNN-pipeline approach.

## 5. Discussion

### 5.1. CNN-Based OIC Performance

In our experiments, we demonstrate that a CNN-based optical registration scheme can achieve similar performance to existing state-of-the-art optical sub-pixel correlators, and in the case of sub-pixel displacements characterized by sharp discontinuities, we exceed it. Reduction of near-field displacement bias in the proximity of sharp discontinuities represents a significant break-through in the characterization of ground displacements using satellite images. Recent applications of OIC in the study of earthquake surface ruptures using high resolution (<1 m) satellite data have tended to focus on (a) characterizing the fault zone width (from the spatial width of the displacement step, e.g. see Figure 7), (b) examining the near-field strain components (e.g., shear and dilatation), to spatially assess the transition between elastic and inelastic deformation (e.g., a yield strength of 0.5% is often assumed), and (c) to quantify the total (on- and off-fault) displacement, which can then be compared with highly localized field-based measurements of slip (assumed to represent purely on-fault displacement), in order to quantify the extent of off-fault displacement. All of these applications require an unbiased correlator, able to accurately and precisely resolve the near-field displacement, which has so far remained elusive (though perhaps not well appreciated by the community). Therefore, potentially all previous studies investigating near-field strain and off-fault deformation based on optical correlation may be biased by to some degree by limitations of the correlation process used to retrieve the displacements. This will be especially apparent when sharp discontinuities become artificially smoothed over a wider zone by the correlation kernel, thereby artificially attributing on-fault slip to the off-fault region. Consequently, estimates of near-field strain will be biased too high, leading to incorrect interpretations about the mechanical behavior of fault zones.

Our approach helps to address this issue, and will lead to more accurate estimates of near-fault displacement in the future; this is the primary value of OIC methods for the study of earthquakes, where the medium and far-field displacements can be more precisely characterized using InSAR techniques.

Additionally, our pipeline, although trained with Landsat-8 satellite imagery, is shown to have general applicability to other optical data sets (e.g., Pleiades). This is likely due to the scale-invariant nature of very small (e.g., 16-by-16 pixel) image windows, which look similar regardless of the sensor used to acquire the data. Nevertheless, our simple pipeline represents a first step toward using data-drive approaches to retrieve sub-pixel displacement maps using optical satellite data.

### 5.2. Implications for Ridgecrest Earthquake Rupture

In the case of the 2019 Ridgecrest earthquake, displacement profiles spanning the fault rupture do not reveal a perfectly sharp displacement. This reflects some degree of smoothing across the discontinuity produced by the sliding correlation window (which considers a larger region of pixels). In addition, we also observe bias/noise related to the inability of a simple translational model to capture a more complex displacement (e.g., a discontinuity) present within a single correlation window.

Recent studies (Antoine et al., 2021; Barnhart et al., 2020; C. Milliner et al., 2021) have made use of traditional correlation approaches using a range of satellite data (WorldView, SPOT6/7, and Pleiades) to constrain the near-field displacement for the Ridgecrest earthquake. Using displacement gradients to compute the 2D strain field, these studies then quantify the fault zone width by measuring the distance over which the fault-perpendicular strain exceeds 0.5% each side of the fault. However, clearly smoothing across a displacement discontinuity will also give a smoothed strain field, thus artificially widening any estimates of fault zone width, and potentially producing incorrect signals of inelastic off-fault deformation. Several factors contribute to this bias, which can be estimated based on the correlation parameters and imagery used. (a) Any displacement estimates within $\pm 0.5$ effective correlation window size of the rupture will be biased. (b) In the case where the fault runs diagonal to the window boundaries (as is the case with Ridgecrest) this distance will be larger (0.707 eff. window size), given the square shape of the correlation window. (c) Any smoothing applied to reduce noise in the displacement map (either non-local means, TVL1, or spatial regularization) will also further widen the discontinuity. (d) When computing the displacement gradients (needed to compute the strain), central difference approaches will further spatially smooth the discontinuity. Finally, downsampling the displacement maps will increase the minimum threshold for which we can estimate fault zone width. Ultimately, the fault zone widths determined from exceedance of any inelastic strain threshold may be biased too large, while the absolute strain values will be biased too low.

In the case of Barnhart et al. (2020), the authors found a fault zone width of ~30 m along-strike, using ampcor (correlation window of $64 \times 64$) and WorldView imagery (0.5 m). However, the smoothing effects of such a correlator on WorldView data will result in a minimum detection limit for fault zone width of 32 m (45 m for a 45°-striking rupture); thus potentially all the slip could have been completely localized, but it would appear as a distributed zone of 30 m wide, thereby leading to incorrect interpretations about the mechanics of slip and off-fault deformation.

In the case of (C. Milliner et al., 2021), the authors document many fault zone widths in the range 10–30 m along-strike, using COSI-Corr (correlation window of $32 \times 32$, stride of 4 pixels, followed by additional non-local means filtering) and SPOT6/7 imagery (1.5 m). The smoothing effects of such a correlator on SPOT6/7 data will result in a minimum detection limit for fault zone width of 24 m (34 m for a 45°-striking rupture). Therefore, it is unclear how reliable such measurements are for inferring details on the mechanics of fault slip, especially in relation to the localization or distribution of slip into the off-fault region.

Finally, in the case of Antoine et al. (2021), the authors correlate 0.5 m Pleaides orthoimages using MicMac. Few details are given regarding the correlation parameters used, so we assume the default settings of $9 \times 9$ window and 0.3 regularization, thus yielding a minimum detection limit of 4.5 m (although it is likely to be somewhat higher, as the spatial regularization results in additional spatial smoothing, as we see in Figure 7). Additional post-processing was applied, to reduce outliers, albeit with few details. The authors downsample their displacement maps (to 2 and 5 m) using a nearest neighborhood approach, before computing the strain. Therefore, the degree of

**Table 4**
*Computation Time of the CNN-Pipeline, COSI-Corr and MicMac: Number of Displacement Estimations (Processed Window Pairs) in 1 s*

| | Number of displacement estimations per sec. | | | |
|---|---|---|---|---|
| | 1 GPU | 32 CPU cores | 8 CPU cores | 1 CPU core |
| Sub-pixel CNN alone | 78.3 $\times 10^3$ | x | x | x |
| 2-step CNN-pipeline | 39.2 $\times 10^3$ | x | x | x |
| COSI-Corr | x | 141 $\times 10^3$ | 42.9 $\times 10^3$ | 6.49 $\times 10^3$ |
| MicMac | x | 23.2 $\times 10^3$ | 23.2 $\times 10^3$ | 5.66 $\times 10^3$ |

*Note*. The window size is similar to previous experiments: 16 × 16 for the sub-pixel CNN, 32 × 32 for the coarse first-step CNN, 15 × 15 for MicMac, 32 × 32 for COSI-Corr, for consistency (see Section 4.1.2).

artificial widening of fault zone width (measured from the 0.5% strain contour) may easily be enhanced by these processing steps, although perhaps not to the degree of Barnhart et al. (2020) or C. Milliner et al. (2021).

The improved performance of our CNN-pipeline in resolving less-biased discontinuities (e.g., Figure 10) could perhaps further reduce any bias associated with the measurements of Antoine et al. (2021) for the Ridgecrest (and other) earthquakes. Future studies will further utilize approaches such as ours to provide more accurate descriptions of near-field displacements needed for addressing how slip is partitioned between the fault plane and the neighboring damage zone.

### 5.3. Computation Time

Our CNN-pipeline is developed within a GPU environment, which offers the potential for significant speed improvements in generating displacement maps from optical data. Our method executes a 1024 × 1024 pixels satellite images pair in 15 s on a single NVIDIA V100 Tensor Core GPU. As COSI-Corr and MicMac run on CPU cores only (Intel Xeon Gold 6342 CPU), we acknowledge the complexity of comparing computation speeds for methods running on either CPU or GPU. On 1024 × 1024 pixels evaluations, using 1 GPU and three different CPU configurations, that is, number of CPU cores (8 is the typical amount of cores available on a powerful laptop), the number displacement estimations (processed window pairs) per second is given in Table 4. First, our 2-step approach is twice as slow as our 1-step sub-pixel model pipeline, as it processes the data twice. Also, MicMac does not seem to gain from using a large number of cores. Our method is already comparable with COSI-Corr in terms of computation time and surpasses MicMac in all configurations, despite our minimal efforts in speed optimization. For example, our CNN approach could be faster with an implementation tuned for GPU-parallel processing (as COSI-Corr makes use of parallel computation on multi-CPU). Furthermore, we currently solve the pixel-scale problem using a similar approach to the sub-pixel refinement, more to demonstrate the potential of these approaches for retrieving complex and large ground displacements. However, the pixel-scale problem is much simpler, and may be better solved using alternative GPU-accelerated disparity estimation techniques (with lower sub-pixel accuracy, but good pixel-scale accuracy, e.g. SGM, (Hernandez-Juarez et al., 2016)). Alternatively, we could instead employ a pyramid-based multi-scale correlation scheme extending from highly downsampled images to the penultimate scale, then upsample the displacement map to seed the sub-pixel refinement stage; this would reduce the computational burden for the pixel-scale correlation step, while also allowing displacements larger than 5 pixels (at the finest scale) to be captured with no additional cost.

We note that a GPU-accelerated implementation of COSI-Corr (or similar frequency-based correlators) will likely give the fastest possible computation times, since the calculation made at each sliding window position is much simpler than with our CNN-model, while also taking advantage of the speed gain from working in the frequency domain. In fact, CNN-based correlation is generally more complex in terms of algorithmic complexity and implementation due to the depth and number of operations, when frequency-based correlation might be simpler but relies heavily on FFT operations. However, such an approach would still be limited by the various assumptions and limitations of traditional local-scale correlation (i.e., homogenous translations within each sliding window), and thus it would not take advantage of the power of data driven approaches to resolve the displacement field with greater accuracy, precision, or superior robustness to noise. Future efforts should explore multi-scale deep learning architectures (e.g., U-Net), which offer the best potential for speed gain and robustness

to high frequency noise; such approaches work with much larger image windows, thereby reducing the computational burden associated with processing many correlation windows. Furthermore, given the deeper architectures and larger image sizes associated with such approaches, they may also benefit from working in the frequency domain, by leveraging the benefits of the FFT for speeding up the convolutional steps.

### 5.4. Limitations

Several limitations remain in our approach, which will be addressed in future studies. First, to create the training data sets, we used the Lanczos resampling algorithm to incorporate sub-pixel shifts, which can ultimately can add sub-pixel bias resulting from high-frequency ringing artifacts. The magnitude of these artifacts varies (in part) as a function of spatial contrasts between image pixels, and which, in turn, may correlate with topography; the magnitude of this bias is typically <1/10th pixel. We also have to keep in mind that when training with synthetic data, we aim to describe a real-world scenario (sliding window crossing a discontinuity) with a simplified representation (see Figure 3 on the right). This representation may lack details in specific configurations, such as multiple discontinuities within the sliding window, high distributed faulting, or a more complex fault pattern, but allows a good generalization and a fast process for the training.

When evaluating the error over larger synthetic displacement maps, we are also subject to interpolation bias introduced by our warping procedure, used to create synthetically displaced satellite images; the magnitude of these resampling errors on the resulting displacement map is generally substantially less than 1/50th pixel. We are also subject to initial mis-registration of the satellite images used to general our training data, which can add bias to our synthetic displacements. We try to mitigate this error by applying a global mis-registration correction to the full pre- and post- images (i.e., not just on the two windows), based on sub-pixel phase correlation, to bring the two images into better alignment; by aligning the two images over a large area, we obtain a more accurate global alignment exceeding 1/10th pixel. Nevertheless, this global alignment step may become less accurate if the illumination conditions between the two images are strongly different.

In addition, we also face the same basic problem of solving for an overly simplistic displacement transform. Although we can now solve for a displacement step function contained within a sliding window, we still ignore higher order terms, rotations, scale changes, etc., which may be important in characterizing the displacement (especially close to sharp discontinuities). Furthermore, insufficient calibration of the CCD sensors, and resampling errors introduced during the ortho-rectification process (prominent in topographically rough regions, and when the DEM is of lower resolution than the image) both introduce sub-pixel bias in the training data. Therefore, the sub-pixel precision of our CNN-approach is likely limited to ~1/10th pixel, reflecting the combined contribution of all these error sources.

### 5.5. Future Developments

First, an upcoming improvement would make our pipeline more flexible by using an external method for the initial pixel-size registration. Additional development of our CNN-based registration approach will likely further enhance performance, particularly in mitigating high frequency noise in the final estimated displacement map. Numerous tests were performed in tuning the model parameters; however, further investigation may lead to improved sub-pixel capability. We also limit the amount of training data to 100k samples (based on sensitivity studies for this parameter, striking a balance between accuracy and computation time); nevertheless, increasing the number of samples may also lead to additional improvements. Our sub-pixel CNN-based model is fixed to window sizes of $16 \times 16$ pixels, thus restricting the level of spatial detail we can resolve; it could be possible to further reduce the window size to obtain additional spatial detail, while relying on multi-scale spatial regularization techniques to mitigate increased noise associated with smaller sliding windows.

The reduction of noise introduced during the generation of training data will also directly enhance the sub-pixel accuracy and precision which we can obtain (and consequently the high frequency spatial noise in the output displacement maps). A significant component of the high frequency noise is likely related to systematic bias resulting from variations in surface reflectance related to differing illumination conditions between the two images. In the case of widespread and freely available optical data sets, such as Landsat-8 or Sentinel-2, where images are acquired at approximately the same time of day (~10a.m. local time), these sources of noise should vary in a predictable manner, thereby enabling this source of bias to be learned and potentially removed using data-driven approaches (Lacroix et al., 2019). On this aspect of topography correlated noise in the images, further

study could incorporate Digital Elevation Models (DEMs) along with the illumination conditions in the displacement estimation process, in the aim of tackling such a source of error. Here we make no effort to apply spatial regularization techniques to smooth high frequency noise; therefore, the fact that our simple data-driven approach already performs on a par with state-of-the-art OIC approaches is encouraging for future development of CNN-based optical registration techniques. One obvious direction for future development will be in the use of multi-scale approaches, such those offered by U-Net architectures, and which have been shown to achieve sub-pixel performance in particular cases (Boukhtache et al., 2021; Ilg et al., 2017). In the case of earthquake, the displacement fields are generally smooth over longer wavelengths, and thus should benefit from multi-scale spatial regularization techniques.

From another perspective, it would be valuable to assess the performance of our model across different natural events, such as landslides and glacier motions. Although beyond the scope of this study, it would be valuable to explore and develop the generalizability of our model, even if these processes are very different from earthquake ruptures. In addition, we only evaluated one real case of Ridgecrest, and we acknowledge the importance of evaluating its performance on other types of faults and geological settings. Future work will include testing the method on real diverse earthquake scenarios.

Finally, a careful benchmarking of our new CNN-based technique against existing techniques will be the subject of a forthcoming paper.

## 6. Conclusion

Our paper presents a complete deep learning framework that extracts ground motion, from pairs of optical satellite images (OIC), with sub-pixel accuracy. By addressing the presence of sharp discontinuities within sliding windows, we propose a new approach that mitigates displacement estimation bias, exceeding the performance of state-of-the-art methods, in the near-field of earthquakes. Our approach therefore reduces the artificial spreading of localized displacement into the near-field region, thus reducing potential estimates of Off-Fault Deformation (made from the difference between localized on-fault slip and accumulated total slip away from the rupture), and Fault Zone Width (e.g., from the fault-normal distance characterized by strains larger than 0.5%; a commonly used threshold assumed to represent the onset of inelastic failure in rocks), compared with displacement maps generated using traditional techniques. We demonstrate this improvement quantitatively using evaluations on high quality synthetic data generated using realistic fault slip models. We also demonstrate that our technique is transferable, achieving state-of-the-art performance on satellite images acquired with different sensors and resolutions; which is not trivial for such data-driven approaches. Several critical challenges must be addressed to advance CNN-based Optical Image Correlation (OIC) for earthquake deformation studies. Sub-pixel bias from the Lanczos resampling algorithm is one limitation of using synthetic data. Initial mis-registration of satellite images introduces additional biases, mitigated but not eliminated by global alignment corrections. Additionally, errors from insufficient CCD sensor calibration and ortho-rectification, especially in topographically rough regions, limit sub-pixel precision. Other sources of errors, such as topography correlated noise, and de-correlation due to a large difference in the acquisition time, are still a great challenge for the estimation of accurate and unbiased displacement fields. Finally, our model's simplistic assumptions overlook rotations, and scale changes, which are likely important features near sharp discontinuities. Future work will focus on leveraging multi-scale deep learning approaches to improve accuracy (including reduction of high frequency noise, and systematic bias), and speed.

**Acknowledgments**
This work is supported by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02). The authors would like to thank Université Grenoble Alpes, INSU PNTS, CNES, and CDP Risk for funding, and GRICAD infrastructure (gricad.univ-grenoble-alpes.fr), which is supported by Grenoble research communities, for the computations. We also thank Solène Antoine for kindly providing the Pleiades orthoimages of the Ridgecrest region.

## Data Availability Statement

The numerical experiments can be reproduced with Pytorch code available (Montagnon, 2024). The generated training databases, as well as the comparisons with state-of-the-art are publicly available (Montagnon, 2023).

## References

Ajorlou, N., Hollingsworth, J., Mousavi, Z., Ghods, A., & Masoumi, Z. (2021). Characterizing near-field surface deformation in the 1990 Rudbar earthquake (Iran) using optical image correlation. *Geochemistry, Geophysics, Geosystems*, *22*(6), e2021GC009704. https://doi.org/10.1029/2021gc009704

Amey, R., Wedmore, L., Gregory, L. C., Hooper, A. J., McCaffrey, K. J., Wilkinson, M. W., et al. (2017). Not just rough around the edges: Fractal properties of exhumed fault surfaces in the Italian Apennines. In *AGU fall meeting abstracts* (Vol. 2017, pp. T52C-08).

Antoine, S. L., Klinger, Y., Delorme, A., Wang, K., Bürgmann, R., & Gold, R. D. (2021). Diffuse deformation and surface faulting distribution from submetric image correlation along the 2019 Ridgecrest, California, ruptures. *Bulletin of the Seismological Society of America*, *111*(5), 2275–2302. https://doi.org/10.1785/0120210036

Avouac, J., & Leprince, S. (2015). Geodetic imaging using optical systems. In *Geodesy* (pp. 387–424). Elsevier Inc.

Barnhart, W. D., Gold, R. D., & Hollingsworth, J. (2020). Localized fault-zone dilatancy and surface inelasticity of the 2019 Ridgecrest earthquakes. *Nature Geoscience*, *13*(10), 699–704. https://doi.org/10.1038/s41561-020-0628-8

Barnhart, W. D., Hayes, G. P., & Gold, R. D. (2019). The July 2019 Ridgecrest, California, earthquake sequence: Kinematics of slip and stressing in cross-fault ruptures. *Geophysical Research Letters*, *46*(21), 11859–11867. https://doi.org/10.1029/2019gl084741

Blanchet, G., Buades, A., Coll, B., Morel, J.-M., & Rougé, B. (2011). Fattening free block matching. *Journal of Mathematical Imaging and Vision*, *41*(1–2), 109–121. https://doi.org/10.1007/s10851-011-0268-0

Boukhtache, S., Abdelouahab, K., Berry, F., Blaysat, B., Grediac, M., & Sur, F. (2021). When deep learning meets digital image correlation. *Optics and Lasers in Engineering*, *136*, 106308. https://doi.org/10.1016/j.optlaseng.2020.106308

Candela, T., Renard, F., Klinger, Y., Mair, K., Schmittbuhl, J., & Brodsky, E. E. (2012). Roughness of fault surfaces over nine decades of length scales. *Journal of Geophysical Research*, *117*(B8), B08409. https://doi.org/10.1029/2011jb009041

Chambolle, A. (2004). An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, *20*(1/2), 89–97. https://doi.org/10.1023/b:jmiv.0000011321.19549.88

Cheng, G., & Barnhart, W. D. (2021). Permanent co-seismic deformation of the 2013 $M_w$7. 7 Baluchistan, Pakistan earthquake from high-resolution surface strain analysis. *Journal of Geophysical Research: Solid Earth*, *126*(3), e2020JB020622. https://doi.org/10.1029/2020jb020622

Dolan, J. F., & Haravitch, B. D. (2014). How well do surface slip measurements track slip at depth in large strike-slip earthquakes? The importance of fault structural maturity in controlling on-fault slip versus off-fault surface deformation. *Earth and Planetary Science Letters*, *388*, 38–47. https://doi.org/10.1016/j.epsl.2013.11.043

Dosovitskiy, A., Fischer, P., Ilg, E., Hausser, P., Hazirbas, C., Golkov, V., et al. (2015). FlowNet: Learning optical flow with convolutional networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2758–2766).

Engwirda, D. (2014). Locally optimal Delaunay-refinement and optimisation-based mesh generation.

Fialko, Y., Sandwell, D., Simons, M., & Rosen, P. (2005). Three-dimensional deformation caused by the Bam, Iran, earthquake and the origin of shallow slip deficit. *Nature*, *435*(7040), 295–299. https://doi.org/10.1038/nature03425

Guizar-Sicairos, M., Thurman, S. T., & Fienup, J. R. (2008). Efficient subpixel image registration algorithms. *Optics Letters*, *33*(2), 156–158. https://doi.org/10.1364/ol.33.000156

Heid, T., & Kääb, A. (2012). Evaluation of existing image matching methods for deriving glacier surface displacements globally from optical satellite imagery. *Remote Sensing of Environment*, *118*, 339–355. https://doi.org/10.1016/j.rse.2011.11.024

Hernandez-Juarez, D., Chacón, A., Espinosa, A., Vázquez, D., Moure, J. C., & López, A. M. (2016). Embedded real-time stereo estimation via semi-global matching on the GPU. *Procedia Computer Science*, *80*, 143–153. https://doi.org/10.1016/j.procs.2016.05.305

Hirschmuller, H. (2007). Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*(2), 328–341. https://doi.org/10.1109/tpami.2007.1166

Hollingsworth, J., Ayoub, F., Doin, M.-P., Daout, S., Perfettini, H., Peltzer, G., & Samsonov, S. (2017). Characterization and removal of shadow bias from optical image correlation; application to the 2013 Balouchistan earthquake. In *AGU fall meeting abstracts* (Vol. 2017, pp. G43A–0909).

Hollingsworth, J., Leprince, S., Ayoub, F., & Avouac, J.-P. (2012). Deformation during the 1975–1984 Krafla rifting crisis, NE Iceland, measured from historical optical imagery. *Journal of Geophysical Research*, *117*(B11), B11407. https://doi.org/10.1029/2012jb009140

Ilg, E., Mayer, N., Saikia, T., Keuper, M., Dosovitskiy, A., & Brox, T. (2017). FlowNet 2.0: Evolution of optical flow estimation with deep networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2462–2470).

Kanade, T., & Okutomi, M. (1994). A stereo matching algorithm with an adaptive window: Theory and experiment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *16*(9), 920–932. https://doi.org/10.1109/34.310690

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*.

Lacroix, P., Araujo, G., Hollingsworth, J., & Taipe, E. (2019). Self-entrainment motion of a slow-moving landslide inferred from Landsat-8 time series. *Journal of Geophysical Research: Earth Surface*, *124*(5), 1201–1216. https://doi.org/10.1029/2018jf004920

Lasserre, C., Peltzer, G., Crampé, F., Klinger, Y., van Der Woerd, J., & Tapponnier, P. (2005). Coseismic deformation of the 2001 $M_w$=7.8 Kokoxili earthquake in Tibet, measured by synthetic aperture radar interferometry. *Journal of Geophysical Research*, *110*(B12), B12408. https://doi.org/10.1029/2004jb003500

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, *1*(4), 541–551. https://doi.org/10.1162/neco.1989.1.4.541

Leprince, S., Ayoub, F., Klinger, Y., & Avouac, J.-P. (2007). Co-registration of optically sensed images and correlation (COSI-Corr): An operational methodology for ground deformation measurements. In *2007 IEEE international geoscience and remote sensing symposium* (pp. 1943–1946).

Leprince, S., Barbot, S., Ayoub, F., & Avouac, J.-P. (2007). Automatic and precise orthorectification, coregistration, and subpixel correlation of satellite images, application to ground deformation measurements. *IEEE Transactions on Geoscience and Remote Sensing*, *45*(6), 1529–1558. https://doi.org/10.1109/tgrs.2006.888937

Leprince, S., Berthier, E., Ayoub, F., Delacourt, C., & Avouac, J.-P. (2008). Monitoring earth surface dynamics with optical imagery. *Eos, Transactions American Geophysical Union*, *89*(1), 1–2. https://doi.org/10.1029/2008eo010001

Lewis, J. P. (1995). Fast template matching. In *Vision interface* (Vol. 95, pp. 15–19).

Magen, Y., Ziv, A., Inbal, A., Baer, G., & Hollingsworth, J. (2020). Fault rerupture during the July 2019 Ridgecrest earthquake pair from joint slip inversion of InSAR, optical imagery, and GPS. *Bulletin of the Seismological Society of America*, *110*(4), 1627–1643. https://doi.org/10.1785/0120200024

Marchandon, M., Hollingsworth, J., & Radiguet, M. (2021). Origin of the shallow slip deficit on a strike slip fault: Influence of elastic structure, topography, data coverage, and noise. *Earth and Planetary Science Letters*, *554*, 116696. https://doi.org/10.1016/j.epsl.2020.116696

Marchandon, M., Wright, T. J., & Hollingsworth, J. (2022). Remote sensing of the earthquake deformation cycle. In *Surface displacement measurement from remote sensing images* (pp. 191–246).

Meade, B. J. (2007). Algorithms for the calculation of exact displacements, strains, and stresses for triangular dislocation elements in a uniform elastic half space. *Computers & Geosciences*, *33*(8), 1064–1075. https://doi.org/10.1016/j.cageo.2006.12.003

Michel, R., & Avouac, J.-P. (2002). Deformation due to the 17 August 1999 Izmit, Turkey, earthquake measured from spot images. *Journal of Geophysical Research*, *107*(B4), ETG-2–ETG2-6. https://doi.org/10.1029/2000jb000102

Milliner, C., Dolan, J., Hollingsworth, J., Leprince, S., & Ayoub, F. (2016). Comparison of coseismic near-field and off-fault surface deformation patterns of the 1992 $M_w$ 7.3 landers and 1999 $M_w$ 7.1 hector mine earthquakes: Implications for controls on the distribution of surface strain. *Geophysical Research Letters*, *43*(19), 10–115. https://doi.org/10.1002/2016gl069841

Milliner, C., & Donnellan, A. (2020). Using daily observations from planet labs satellite imagery to separate the surface deformation between the 4 July $M_w$ 6.4 foreshock and 5 July $M_w$ 7.1 mainshock during the 2019 Ridgecrest earthquake sequence. *Seismological Research Letters*, *91*(4), 1986–1997. https://doi.org/10.1785/0220190271

Milliner, C., Donnellan, A., Aati, S., Avouac, J.-P., Zinke, R., Dolan, J. F., et al. (2021). Bookshelf kinematics and the effect of dilatation on fault zone inelastic deformation: Examples from optical image correlation measurements of the 2019 Ridgecrest earthquake sequence. *Journal of Geophysical Research: Solid Earth*, *126*(3), e2020JB020551. https://doi.org/10.1029/2020jb020551

Milliner, C., Sammis, C., Allam, A., Dolan, J., Hollingsworth, J., Leprince, S., & Ayoub, F. (2016). Resolving fine-scale heterogeneity of co-seismic slip and the relation to fault structure. *Scientific Reports*, *6*(1), 27201. https://doi.org/10.1038/srep27201

Milliner, C. W., Dolan, J. F., Hollingsworth, J., Leprince, S., Ayoub, F., & Sammis, C. G. (2015). Quantifying near-field and off-fault deformation patterns of the 1992 $M_w$ 7.3 l Anders earthquake. *Geochemistry, Geophysics, Geosystems*, *16*(5), 1577–1598. https://doi.org/10.1002/2014gc005693

Montagnon, T. (2023). Synthetic datasets and evaluations for sub-pixel displacements estimation from optical satellite images with deep learning [Dataset]. *Recherche Data Gouv*. https://doi.org/10.57745/UOGRPY

Montagnon, T. (2024). cnn4l-discontinuities [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.12170087

Montagnon, T., Hollingsworth, J., Pathier, E., Marchandon, M., Dalla Mura, M., & Giffard-Roisin, S. (2022). Sub-pixel optical satellite image registration for ground deformation using deep learning. In *2022 IEEE international conference on image processing (ICIP)* (pp. 2716–2720).

Nefian, A. V., Husmann, K., Broxton, M., To, V., Lundy, M., & Hancher, M. D. (2009). A Bayesian formulation for sub-pixel refinement in stereo orbital imagery. In *2009 16th IEEE international conference on image processing (ICIP)* (pp. 2361–2364).

Nikkhoo, M., & Walter, T. R. (2015). Triangular dislocation: An analytical, artefact-free solution. *Geophysical Journal International*, *201*(2), 1119–1141. https://doi.org/10.1093/gji/ggv035

Padilla, A. M. R., & Oskin, M. E. (2023). A probabilistic displacement hazard assessment framework for distributed ruptures from strike-slip earthquakes.

Page, M. T., Custódio, S., Archuleta, R. J., & Carlson, J. (2009). Constraining earthquake source inversions with GPS data: 1. Resolution-based removal of artifacts. *Journal of Geophysical Research*, *114*(B1), B01314. https://doi.org/10.1029/2007jb005449

Pan, B., Qian, K., Xie, H., & Asundi, A. (2009). Two-dimensional digital image correlation for in-plane displacement and strain measurement: A review. *Measurement Science and Technology*, *20*(6), 062001. https://doi.org/10.1088/0957-0233/20/6/062001

Pinel-Puysségur, B., De Zan, F., & Champenois, J. (2022). The interferometric phase: Unwrapping and closure phase. In *Surface displacement measurement from remote sensing images* (pp. 155–189).

Ponti, D., Blair, J., Rosa, C., Thomas, K., Pickering, A., Morelan, A., & Dawson, T. (2020). *Digital datasets documenting surface fault rupture and ground deformation features produced by the Ridgecrest M6. 4 and M7. 1 earthquake sequence of July 4 and 5, 2019*. US Geological Survey.

Rosen, P. A., Hensley, S., Peltzer, G., & Simons, M. (2004). Updated repeat orbit interferometry package released. *Eos, Transactions American Geophysical Union*, *85*(5), 47. https://doi.org/10.1029/2004eo050004

Ross, Z. E., Idini, B., Jia, Z., Stephenson, O. L., Zhong, M., Wang, X., et al. (2019). Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence. *Science*, *366*(6463), 346–351. https://doi.org/10.1126/science.aaz0109

Rosu, A.-M., Pierrot-Deseilligny, M., Delorme, A., Binet, R., & Klinger, Y. (2015). Measurement of ground displacement from optical satellite image correlation using the free open-source software MicMac. *ISPRS Journal of Photogrammetry and Remote Sensing*, *100*, 48–59. https://doi.org/10.1016/j.isprsjprs.2014.03.002

Rupnik, E., Daakir, M., & Pierrot Deseilligny, M. (2017). MicMac–a free, open-source solution for photogrammetry. *Open Geospatial Data, Software and Standards*, *2*(1), 1–9. https://doi.org/10.1186/s40965-017-0027-2

Scambos, T. A., Dutkiewicz, M. J., Wilson, J. C., & Bindschadler, R. A. (1992). Application of image cross-correlation to the measurement of glacier velocity using satellite image data. *Remote Sensing of Environment*, *42*(3), 177–186. https://doi.org/10.1016/0034-4257(92)90101-o

Scholz, C. H. (1982). Scaling laws for large earthquakes: Consequences for physical models. *Bulletin of the Seismological Society of America*, *72*(1), 1–14.

Scott, C. P., Arrowsmith, J. R., Nissen, E., Lajoie, L., Maruyama, T., & Chiba, T. (2018). The M7 2016 Kumamoto, Japan, earthquake: 3-D deformation along the fault and within the damage zone constrained from differential lidar topography. *Journal of Geophysical Research: Solid Earth*, *123*(7), 6138–6155. https://doi.org/10.1029/2018jb015581

Simonovsky, M., Gutiérrez-Becker, B., Mateus, D., Navab, N., & Komodakis, N. (2016). A deep metric for multimodal registration. In *International conference on medical image computing and computer-assisted intervention* (pp. 10–18).

Sutton, M. A., Orteu, J. J., & Schreier, H. (2009). *Image correlation for shape, motion and deformation measurements: Basic concepts, theory and applications*. Springer Science & Business Media.

Tong, X., Ye, Z., Xu, Y., Gao, S., Xie, H., Du, Q., et al. (2019). Image registration with Fourier-based image correlation: A comprehensive review of developments and applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(10), 4062–4081. https://doi.org/10.1109/jstars.2019.2937690

Tong, X., Ye, Z., Xu, Y., Liu, S., Li, L., Xie, H., & Li, T. (2015). A novel subpixel phase correlation method using singular value decomposition and unified random sample consensus. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(8), 4143–4156. https://doi.org/10.1109/tgrs.2015.2391999

Van Puymbroeck, N., Michel, R., Binet, R., Avouac, J.-P., & Taboury, J. (2000). Measuring earthquakes from optical satellite images. *Applied Optics*, *39*(20), 3486–3494. https://doi.org/10.1364/ao.39.003486

Wang, S., Quan, D., Liang, X., Ning, M., Guo, Y., & Jiao, L. (2018). A deep learning framework for remote sensing image registration. *ISPRS Journal of Photogrammetry and Remote Sensing*, *145*, 148–164. https://doi.org/10.1016/j.isprsjprs.2017.12.012

Wiemker, R. (1996). Registration of airborne scanner imagery using akima local quintic polynomial interpolation. In *Presented at the second international airborne remote sensing conference and exhibition* (Vol. 24, p. 27).

Yang, R., Li, Y., Zeng, D., & Guo, P. (2022). Deep DIC: Deep learning-based digital image correlation for end-to-end displacement and strain measurement. *Journal of Materials Processing Technology*, *302*, 117474. https://doi.org/10.1016/j.jmatprotec.2021.117474

Yang, X., Kwitt, R., Styner, M., & Niethammer, M. (2017). Quicksilver: Fast predictive image registration–a deep learning approach. *NeuroImage*, *158*, 378–396. https://doi.org/10.1016/j.neuroimage.2017.07.008

Ye, F., Su, Y., Xiao, H., Zhao, X., & Min, W. (2018). Remote sensing image registration using convolutional neural network features. *IEEE Geoscience and Remote Sensing Letters*, *15*(2), 232–236. https://doi.org/10.1109/lgrs.2017.2781741

Zach, C., Pock, T., & Bischof, H. (2007). A duality based approach for realtime TV-$L^1$ optical flow. In *Pattern recognition: 29th DAGM symposium, Heidelberg, Germany, September 12-14, 2007* (Vol. 29, pp. 214–223). https://doi.org/10.1007/978-3-540-74936-3_22

Zinke, R., Hollingsworth, J., Dolan, J. F., & Van Dissen, R. (2019). Three-dimensional surface deformation in the 2016 $M_w$ 7.8 Kaikōura, New Zealand, earthquake from optical image correlation: Implications for strain localization and long-term evolution of the Pacific-Australian plate boundary. *Geochemistry, Geophysics, Geosystems*, *20*(3), 1609–1628. https://doi.org/10.1029/2018gc007951