Geophysical Research Letters

Supporting Information for

Prediction of Off-Fault Deformation from Experimental Strike-slip Fault Structures using the Convolutional Neural Networks

L. Chaipornkaew[1], H. Elston[2], M. Cooke[2], T. Mukerji[1], S.A. Graham[1]

[1]Stanford University, [2]University of Massachusetts Amherst

Contents of this file

- Texts S1 to S5
- Figures S1 to S2
- Tables S1 to S3

Introduction

Please find supplementary texts, figures, and tables below to explain our experiment setups, data preparation, methodology, and findings.

Text S1. Kaolin

We use a wet kaolin tile #6 pottery clay to examine echelon fault growth and linkage because this material simulates several critical elements of crustal faulting. Wet kaolin displays localized and slickenlined fault surfaces (Henza et al., 2010) that allow for tracking of fault linkage and evolution. The wet kaolin has non-zero cohesion, which facilitates continued slip along fault surfaces that are no longer preferentially oriented, consistent with observations of crustal faulting. At failure, the wet kaolin shows rate and state behavior and normal stress dependency consistent with crustal materials (Cooke & van der Elst, 2012). Finally, the wet kaolin is a viscoelastic material that shows off-fault stress relaxation that can simulate pervasive off-fault permanent deformation in the crust. This last property allows us to investigate how the kinematic

efficiency of strike-slip faults changes as faults evolve within our experimental apparatus.

The properties of the wet kaolin depend on water content (Eisenstadt & Sims, 2005). We adjust the water content of the wet kaolin so that fall cone tests report 100-107 Pa undrained shear strength (DeGroot & Lunne, 2007; Table S1). The resulting water content is ~ 71-77% by weight for these experiments. This target strength ensures that 1cm within the experiment scales as 1 to 2 km of crust. The range reflects the range of estimates for km-scale strength of rock material in the upper few kms of the crust. We follow procedures described in Hatem et al. (2017) for preparation of the wet kaolin.

Text S2. Experimental Setup

Stepper motors move one side of a split box to create basal strike-slip strain that is either localized or distributed via an elastic sheet attached to the base of both boxes. To prevent rippling of the elastic upon shear we pre-stretch the 1.5 cm wide elastic sheet to a width of 2.5 cm. Throughout the experiments with the elastic shear, the basal shear is distributed over 2.5 cm. As the localized or distributed basal shear drives upward propagation of the fault, photos taken every ~0.3 mm of plate displacement document the evolution of the fault at the top surface of the clay pack. Red and black sand grains distributed on the surface of the clay allow us to calculate the incremental horizontal displacement fields from the photos of the clay surface using Digital Image Correlation (DIC) techniques. We use the matlab based PIVlab for the DIC techniques following techniques described in Toeneboehn et al. (2018). The distribution of sand and timing of photos allows for <0.01 mm uncertainty of horizontal displacement between successive images of our experiment measured at points with spacing of 1 mm.

Nine experiments have distributed basal shear with an elastic sheet and 7 experiments have localized basal shear. The experiments span a range of applied plate velocities: 0.25 mm/min to 1.5 mm/min (Table S1).

Strain rate maps derived from the incremental horizontal displacement fields allow us to delineate active faults at all stages of the experiment and track the evolution of the early echelon fault arrays. The total strain rate is the sum of the incremental vorticity and divergence maps calculated from gradients in the incremental displacement field. Because the divergence rate is very low in these strike-slip experiments, it contributes considerable noise to the total strain rate field. Consequently, here we only consider vorticity rate fields when mapping faults. We identify faults from the vorticity maps using several stages of filtering. First an adaptive image threshold filter (Bradley & Roth, 2007) with 10% sensitivity distinguishes regions that have shear strain 10% above the local average strain. This

filter picks out localized faults within the region of distributed shear strain that develops above the transition between the basal plates without having to prescribe a global shear strain threshold that depends on motor speed and distributed nature of basal shear. This method also detects faults when background strain varies with position in the experiment. However, the low neighborhood shear strain in regions far from basal plate transition means that noise within the strain signal exceeds 10% of the local average strain. For this reason, we also exclude areas with vorticity rate below a background of 0.025 times the incremental displacement of the basal plate. For these experiments with typical incremental plate displacement between photographs of 0.3 mm, regions with shear strain less than 0.0075 are excluded. This is nearly an order of magnitude less than the shear strain along the faults (Fig. S1). Finally, to further reduce noise within the fault maps, we exclude remaining regions of high strain with fewer than 8 connected data points, corresponding to less than 9 mm2 total area. We use the fault maps to find the amount of displacement along the faults compared to the applied displacement. The kinematic efficiency is the ratio of fault slip to total displacement, which serves as the labels to the CNN.

Text S3. Data Pre-Processing

We set up a pipeline to pre-process raw experimental dataset for training CNNs. The pre-processing, such as subsampling the whole fault map into five cropped slices, expands the dataset by 5 folds for raw data. Then, we apply common geometric transformation (Flip, Zoom, Shift) via the ImageDataGenerator class of Keras to significantly manifold both size and variation of the dataset. While 'Flip' is a binary augmentation. Random 'Zoom' and 'Shift' allow all possible floats within the specified intervals to be randomly applied during training. See Figure S2.

Since each experiment variation is repeated, we sub-divide the complete dataset into three statistically equivalent subsets for training, evaluation, and test purposes. Table S1 displays naming schemes between experimental dataset (by dates) and dataset referred to in GitHub (by experiment variations).

Text S4. CNN's hyperparameters selection and performance.

We search and attain a combination of hyperparameters that achieve the best performance in predicting kinematic efficiency from fault maps. This combination can successfully combat noisy validation loss and poor model performance (<70% accuracy) during early search. Since the number of kernels indicate the complexity of our CNN, we selected the smallest, most efficient network that performs well. Learning rate and batch size are two of the most sensitive hyperparameters for our models while momentum and dropout rate are less sensitive. However, since each hyperparameter has interactive effects to model performance, we propose the combination of parameters reported in Table S2 to be most optimal for our dataset.

We repeat training sessions using the final combination of hyperparameters but varying randomized initialization of the network weights. Selecting out of the five trials, the one that has the best performance is used for prediction tasks in an unseen dataset (Table S3). Since our model starts to over fit beyond Epoch 50, we impose early stopping. We choose the best CNN model, which produces the lowest loss in the evaluation dataset, also highest in mini-batch accuracy for both training and evaluation datasets.

Mini-batch accuracy during training (i.e., Table S3 and Fig 2c) is a performance of the model on the mini-batch of only 256 samples. On the other hand, model accuracy for prediction task (i.e., Fig. 3) is an overall accuracy of the whole dataset.

Text S5. Crustal Fault Maps preparation for CNN prediction

We create windows across the fault traces from the three crustal faulting studies that are equivalent in dimension (128x64 pixels) and scale to the experimental windows used to train the CNN. In order to have more than one fault window for each crustal study, we use the minimum range of wet kaolin length scaling, where 1 cm of clay scales to 1 km of crust. This scaling provides 2-4 windows for each study. The 6.4 cm wide experimental fault windows then scale to 6.4 km wide crustal fault windows. The thickness of the fault traces (i.e., number of pixels) may impact the CNN prediction of kinematic efficiency. Active fault traces within the experimental data range from 2 to 4 pixels. Fault zones may thin with system maturity (Hatem et al., 2017). To avoid presuming a fault maturity and fault zone width for the crustal faults, we use crustal fault maps with thicknesses of 2, 3, and 4 pixels. To be consistent with the experimental dataset that trained the CNN, all windows are oriented so that the maximum shear strain direction for the region parallels the short edge of the window. For the San Andreas fault study, we used the max shear direction from SCEC's community geodetic model horizontal velocities while the displacement analysis of Scott et al. (2018) showed the max shear direction. For each window along the crustal fault maps, the CNN predicts the kinematic efficiency directly from the fault trace geometry without having to know any information about the deformation field.

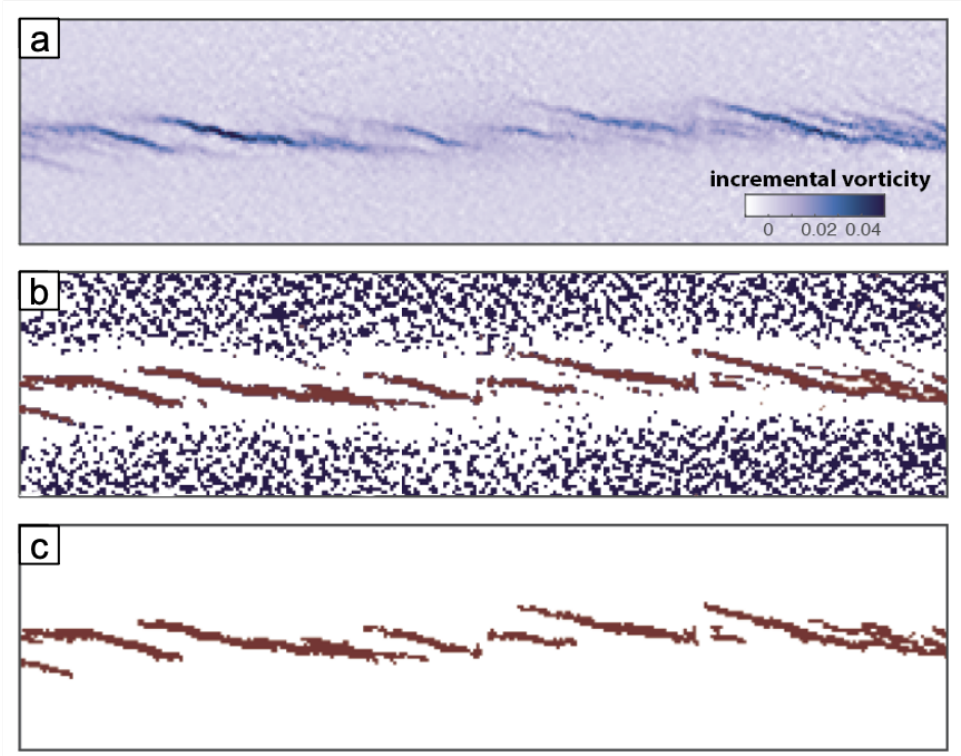

Figure S1. (a) Incremental vorticity map of early echelon faults within experiment 5_16_19, in which clay deforms over an elastic sheet. (b) The adaptive threshold highlights regions of strain that are greater than the local background. Both blue and red regions of local high strain are detected by the adaptive threshold, but we eliminate the blue regions by setting a lower shear strain bound. (c) Further filtering removes small faults, and the resulting map is used as the active fault pattern at that stage of the experiment.

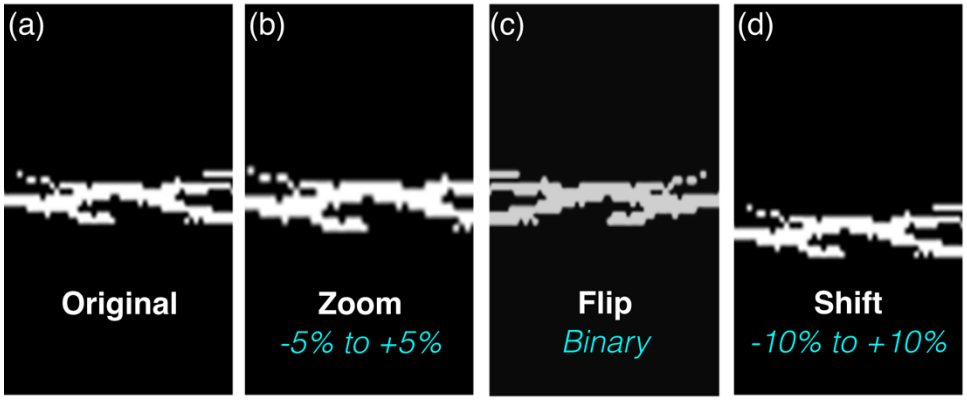

Figure S2. a) The original sliced fault map. b)-d) A display of one realization for each transformation: Zoom, Flip, Shift. The Keras ImageDataGenerator transforms raw data randomly to feed each training epoch any possible realization within the specified intervals.

Distributed Basal Shear				
Experiment Name (lab)	**Loading (mm/min)**	**Initial Clay strength (Pa)**	**Experiment name (GitHub)**	**Data Split**
5_16_19	0.25	100	EB_025_1	Train
5_18_19	0.25	106	EB_025_2	Train/Eval/Test
5_23_19	0.25	102	EB_025_3	Train/Eval/Test
5_17_19	0.5	104	EB_050_1	Train
5_20_19	0.5	105	EB_050_2	Train/Eval/Test
5_22_19	0.5	101	EB_050_3	Train/Eval/Test
5_19_19	1.5	104	EB_150_1	Train
5_24_19	1.5	102	EB_150_2	Train/Eval/Test
5_25_19	1.5	102	EB_150_3	Train/Eval/Test
Localized Basal Shear				
Experiment Name (lab)	**Loading (mm/min)**	**Initial Clay strength (Pa)**	**Experiment name (GitHub)**	**Data Split**
6_07_19	0.25	104	PP_025_1	Train
6_18_19	0.25	102	PP_025_2	Train/Eval/Test
6_19_19	0.5	106	PP_050_1	Train
6_24_19	0.5	106	PP_050_2	Train/Eval/Test
6_04_19	1.0	107	PP_100_1	Train
6_13_19	1.0	104	PP_100_2	Train/Eval/Test
6_03_19	1.5	107	PP_150_1	Train/Eval/Test

Table S1. Experimental Variations, including loading conditions (mm/min) and initial clay strength (Pa). Since each experiment set-up is repeated, we purposefully split the data into training, evaluation, and testing.

Hyperparameters	**Search Ranges**	**Selected Values**
# Kernels in the first CNN layer	2 – 8	4
Learning rate	1e-5 – 1e-2	5e-3
Batch size	32 – 512	256
Momentum	0.75 – 0.95	0.8
Dropout Rate	10% - 50%	20%

Table S2. Final hyperparameters selection

Training Sessions	Best Epoch #	Lowest Eval Loss	Mini-Batch Accuracy	
			Train Dataset	Eval Dataset
1	37	2.621	90.5%	94.2%
2	46	2.403	91.8%	95.8%
3	45	2.492	92.4%	94.7%
4	77	2.120	93.6%	96.1%
5	41	2.138	92.9%	96.6%

Table S3: Performance report for repeated CNN training sessions using the finalized combination of hyperparameters and randomized initialization. The entries of this table are ranked by their accuracy performance of the evaluation dataset. The 'best model' (entry #5) is selected for prediction tasks of an unseen dataset of this paper. Model performance generally deteriorate after epoch 50. However, entry #4 show an out-of-trend performance with exceptionally high training accuracy, though the evaluation accuracy is still lower than the selected 'best model'.

Reference:

Bradley, D., & Roth, G. (2007). Adaptive thresholding using the integral image. Journal of Graphics Tools, 12(2), 13–21.

Cooke, M. L., & van der Elst, N. J. (2012). Rheologic testing of wet kaolin reveals frictional and bi-viscous behavior typical of crustal materials. Geophysical Research Letters, 39(1).

DeGroot, D., & Lunne, T. (2007). Measurement of Remoulded Shear Strength. Norwegian Geotechnical Institute. Report, 20061023–1.

Eisenstadt, G., & Sims, D. (2005). Evaluating sand and clay models: do rheological differences matter? Journal of Structural Geology, 27(8), 1399–1412.

Hatem, A. E., Cooke, M. L., & Toeneboehn, K. (2017). Strain localization and evolving kinematic efficiency of initiating strike-slip faults within wet kaolin experiments. Journal of Structural Geology. https://doi.org/10.1016/j.jsg.2017.06.011

Toeneboehn, K., Cooke, M. L., Bemis, S. P., Fendick, A. M., & Benowitz, J. (2018). Stereovision combined with particle tracking velocimetry reveals advection and uplift within a restraining bend simulating the Denali fault. Frontiers in Earth Science, 6, 152.