Biophysical Journal
Article

Biophysical Society

Traction force microscopy by deep learning

Yu-li Wang[1,*] and Yun-Chu Lin[1]
[1]Department of Biomedical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania

ABSTRACT Cells interact mechanically with their surroundings by exerting and sensing forces. Traction force microscopy (TFM), purported to map cell-generated forces or stresses, represents an important tool that has powered the rapid advances in mechanobiology. However, to solve the ill-posed mathematical problem, conventional TFM involved compromises in accuracy and/or resolution. Here, we applied neural network-based deep learning as an alternative approach for TFM. We modified a neural network designed for image processing to predict the vector field of stress from displacements. Furthermore, we adapted a mathematical model for cell migration to generate large sets of simulated stresses and displacements for training and testing the neural network. We found that deep learning-based TFM yielded results that resemble those using conventional TFM but at a higher accuracy than several conventional implementations tested. In addition, a trained neural network is appliable to a wide range of conditions, including cell size, shape, substrate stiffness, and traction output. The performance of deep learning-based TFM makes it an appealing alternative to conventional methods for characterizing mechanical interactions between adherent cells and the environment.

SIGNIFICANCE Traction force microscopy has served as a fundamental driving force for mechanobiology. However, its nature as an ill-posed inverse problem has raised serious challenges for conventional mathematical approaches. This study, facilitated by large sets of simulated data, describes an alternative approach using deep learning for the calculation of cell-generated stresses. By adapting a neural network architecture widely used for image processing for processing mechanical activities, we show that deep learning is able to predict traction stresses with speed, accuracy, resolution, and versatility.

INTRODUCTION

Mechanobiology examines the input and output of mechanical forces by cells. Contractile forces generated by the actin-myosin cytoskeleton are transmitted to the exterior environment via transmembrane proteins such as integrins (1). Referred to as traction forces, these forces are located predominantly near the cell periphery and directed toward the cell center (2). Active forces are located primarily at maturing focal adhesions near the front, whereas resistive forces are located at mature focal adhesions near the rear (3,4).

In addition to propelling cell migration, traction forces are believed to perform important functions such as organizing the extracellular matrix (5), probing mechanical properties of the environment (6), and sensing the state of the cell itself such as shape, size, and state of migration (7,8). Adhesive cells are also keenly sensitive to mechanical forces transmitted via the surrounding matrix (9), fluid shear (10), or cell-cell contact (11). These mechanical signals elicit many profound responses to affect cell migration (12), growth (13), and differentiation (14).

Methods for mapping traction forces or stresses, referred to as traction force microscopy (TFM (2,15–17)), represent a fundamental tool for mechanobiology. TFM is commonly performed by culturing cells on a substrate of elastic material such as polyacrylamide, embedded with particle markers for mapping force-induced displacements. Substrate displacements then serve as the input for calculating the distribution of forces or stresses based on linear elastic theory.

TFM based on elastic substrates such as polyacrylamide is typically challenged by its mathematical nature as an ill-posed inverse problem, in which a unique solution may not exist and the results are prone to artifacts if one simply tries to minimize the goodness of fit between predicted and measured displacements (18). To mitigate the problems, most conventional TFM approaches include a regularization term in the fitting function to apply penalties to both deviations from targeted values and complexity of the answer. As

Submitted January 4, 2021, and accepted for publication June 8, 2021.
*Correspondence: yuliwang@andrew.cmu.edu
Editor: Catherine Galbraith.
https://doi.org/10.1016/j.bpj.2021.06.011

a result, it sacrifices some accuracy in exchange for reduced artifacts or unrealistic features (2,18). Selecting the algorithm and setting the parameter for regularization prove nontrivial and require a delicate balance between high accuracy and minimal artifacts (19). A related challenge is that both the forward calculation of displacements from stress and inverse calculation of stress from displacements are affected by singular elements in Green's tensor that describes their relationships.

Neural network-based deep learning has been deployed as a powerful method for solving ill-posed problems (20). It involves the optimization of a pipeline of convolution/correlation operations with the goal of transforming the input (e.g., substrate displacements) into targeted solution (e.g., the distribution of stresses (21)). As an approach fundamentally different from conventional TFM, it avoids the challenge of singularity during the prediction and typically invokes regularization only during training while avoiding regularization altogether during prediction.

To deploy deep learning for TFM (DL-TFM), we have adapted a widely used neural network architecture, U-Net, designed originally for multiresolution segmentation of biomedical images (22). The network consists of a single end-to-end pipeline of encoding and decoding operations to preserve features of different sizes. Several skip connections serve to concatenate the information at different stages of the pipeline to balance the performance of global location versus local context with a high efficiency (23). The success of U-Net has prompted its applications in a wide range of computer vision tasks beyond its original purpose.

Although a network architecture derived from U-Net has been reported for TFM (24), the training process relied upon a limited set of experimental data processed by convention TFM, which thereby carried over the errors and limitations of conventional TFM. To minimize the dependence on conventional TFM and to provide large and precise data sets for machine learning, we have applied a simulation model to generate stresses and displacements that mimic the pattern seen in cultured cells. In addition, we have replaced two-dimensional (2D) convolutions in U-Net with three-dimensional (3D) convolutions to reproduce the relationship between stress and displacement fields according to the linear elastic theory. We hereby report the performance of DL-TFM with respect to its speed, accuracy, resolution, and scalability, using several open-access implementations of conventional TFM known as Fourier transform traction cytometry (FTTC (15,25–27)) as the reference.

MATERIALS AND METHODS

Generation of simulated data for training and testing the neural network

Simulated stress fields were generated by modifying our published model for cell migration (28). This model has several major advantages for this purpose. First, its parameters describe local signals for protrusion or retraction, which are known to bear some correlation with traction forces and may be reasonably used for generating a hypothetical stress distribution. Second, it provides information around the entire cell circumstance, which allowed us to simulate a complete stress field underneath a cell. Third, its high versatility allows the generation of a wide variety of realistic-looking cell shapes (28), which provided us with a large and diverse set of data for training and testing.

The model implemented the mechanism of local excitation-global inhibition (29) for controlling the protrusion and retraction activities. Protrusion signals were local, represented by a value specific to each location around the cell periphery, whereas retraction signals were global, represented by a single location-independent value. The periphery extended outward if the local excitation signal exceeded the global retraction signal and retracted inward if the opposite was true. The global retraction signal was proportional to the total protrusion signals around the cell under the assumption that local protrusions generate retraction signals that propagated rapidly as a negative feedback to suppress competing protrusive activities elsewhere. Protrusion signals, on the other hand, were generated locally and stochastically with the level governed by diffusion and decay. A positive feedback loop increased the probability of its generation at locations where the level was high. Implementation of these simple rules led to a highly versatile "automaton" capable of creating a wide range of shapes mimicking various cell types (28). In addition, simultaneous extension and retraction at opposite ends led to migration with the characteristics of persistent random walk.

Because traction forces were known to colocalize with focal adhesions, near the front for active propulsion and near the rear for passive resistance (3), we assumed that the magnitude of traction stress was proportional to the absolute value of net signals, calculated as protrusion signal subtracted by retraction signal, such that strong positive signals to extend at the front or negative signals to retract at the rear were correlated with strong traction stresses. Stress vectors were placed at a random distance from cell periphery following Poisson distribution and in a direction toward the cell center. Simulated cells whose stresses were balanced to $<10\%$ were selected for training or testing.

Substrate displacements were calculated from traction stresses based on a Poisson ratio of 0.45, using Boussinesq's equations for deriving the Green's tensor (2), which was applied via a set of 2D convolution operations followed by elementwise additions to generate displacement vectors. The central element of Green's tensor, which in theory equals infinity due to its inverse dependence on the distance from the source of stress (15), was estimated empirically based on minimizing the error of FTTC (15). This was justified for the current study because the performance of DL-TFM was to be compared against that of FTTC. We assigned a range of testing values to the central element for generating the corresponding displacement fields from a known stress field. FTTC was then applied to the displacement fields to obtain calculated stress fields. The central value that yielded the smallest normalized root mean-squared error (RMSE) between calculated stress field and the known ground truth was used for calculating the displacement for training and testing DL-TFM.

Three similar sets of simulated cells were generated for training neural networks for field sizes of 104×104, 160×160, and 256×256 pixels. Displacements were generated using a Poisson ratio of 0.45 and substrate Young's modulus of 10, 6.5, and 4.1 N/pixel for the respective field size. The decreasing Young's moduli were designed to generate a similar size of displacement relative to the size of the field. Stress vectors were normalized for an average magnitude of 1 N/pixel within the cell. From the stress field of 160×160 pixels, the corresponding stress fields for other sizes were generated by nearest-neighbor interpolation before recalculating the respective displacements. Predicted stress values, as provided by the activations of the output regression layer of the neural network, were scaled by a factor that equals Young's modulus of the substrate in Pascals divided by the Young's modulus for the training set in Newtons/pixel^2 to obtain the predicted stress field in Pascals.

For testing the performance of trained neural networks, simulated cells were generated in the same manner as for the training set but with a

maximal stress magnitude of 6000 Pa on a substrate with a Young's modulus of 10,670 Pa, which represented typical values in our experiments with NIH 3T3 cells. Simulated noise of displacements was created using a Gaussian random number generator. The magnitude of the noise was measured from experimental displacement fields by calculating the SD of displacements within a 20 × 20-pixel moving window that scanned across the entire field including regions far away from the cell. The minimal SD was used as the magnitude of noise.

TFM

MATLAB (The MathWorks, Natick, MA) version 2020b was used for both FTTC and DL-TFM. Three implementations of FTTC were tested, an ImageJ plugin (26) and two open-source MATLAB programs ((15,27) https://github.com/DanuserLab/TFM/tree/master/software and https://github.com/CellMicroMechanics/Easy-to-use_TFM_package). Because the two MATLAB programs showed a similar accuracy better than the ImageJ plugin, the MATLAB program named reg_fourier_TFM.m in https://github.com/DanuserLab/TFM/tree/master/software was used for generating the FTTC data in this study.

Displacements and stresses were represented as 2D vectors, which must be represented by two numbers at each location. Therefore, we applied 3D tensors to represent these vector fields, using two separate planes along the z axis for holding x and y components. The tensors were of a size of S × S × 2, where S ∈ {104, 160, 256} specified the size of a square field, whereas the third dimension held x and y components of the stress or displacement vectors. U-Net was constructed layer by layer from scratch, using 3D convolution for all the convolution operations (Fig. S1). 3D convolution proved to be essential for TFM because the generation of displacements from stresses, as described by Boussinesq's equations, involved linear combinations of stresses along both x and y directions, placed on separate planes of the z axis. In contrast, 2D convolution as applied typically to independent color channels creates only linear combinations within the same plane. Testing with 2D convolution showed that it failed to reach the same accuracy as yielded by 3D convolution despite the convergence during training.

Training was conducted using the optimization algorithm of stochastic gradient descent with momentum for 50–100 epochs until the RMSE reached below three. Minibatch size was 30 for images of 104 × 104, 20 for 160 × 160, and 13 for 256 × 256; the reduction in size was compelled by the limited memory size of the graphics processing unit. Proper training rates were important for the balance between the stability of convergence and the speed of learning. The training rate was set initially at 6×10^{-4} to decrease by a factor of 0.7943 every 10 epochs such that the rate dropped by 10 folds over 100 epochs. The L2 regularization factor for training, designed to prevent overfitting by balancing the goodness of fit against the complexity of kernel weights, was set at 5×10^{-4}. Other hyperparameters were left as default values implemented in MATLAB, including the forward loss function based on half mean-squared error.

Using a Dell Precision E5450 laptop computer (Round Rock, TX) with an Intel Xeon E-2176M CPU (Santa Clara, CA) running at 2.70 GHz, 32 GB of random-access memory, and an NVIDA Quadro P2000 GPU (Santa Clara, CA), the training process took up to 24 h for fields of 256 × 256 pixels and 7.5 h for fields of 104 × 104 pixels (Table S1). MATLAB codes for constructing, training, and deploying the U-Net neural network have been posted on GitHub (https://github.com/yuliwang8177/DL-TFM). Training and testing data sets were disseminated separately via Box due to the large size (https://cmu.box.com/s/n34hbfopwa3r6rftvtfn4ckc403hk43d).

Substrate preparation and cell culture

Polyacrylamide substrates were prepared by mixing 8% acrylamide (Bio-Rad Laboratories, Hercules, CA), 0.1% bis-acrylamide (Bio-Rad Laboratories), and 0.2 μm fluorescent beads at a dilution of 1:1000 (Molecular Probes, Carlsbad, CA). After adding a photoinitiator, lithium phenyl-2,4,6-trimethylbenzoylphosphinate (Allevi, Philadelphia, PA) at a concentration of 1 mg/mL, an aliquot of 25 μL was pipetted onto a large #1 coverslip pretreated with 0.3% (v/v) bind-silane in 95% ethanol (Cytiva, Marlborough, MA) for bonding with polyacrylamide. A coverslip 22 mm in diameter coated with activated gelatin was placed on top to allow the transfer of gelatin to the surface of thin gel (7). The mixture was exposed to 365 nm UV for 3 min under a tabletop UV lamp with two 15 W tubes (UVP, Jena, Germany). The top coverslip was then gently removed with a pair of fine tweezers. Young's modulus of the gel was measured as 10,670 Pa using a custom-designed light-weight flat punch. The substrate was sterilized with 260 nm UV for 20 min before use. NIH 3T3 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% donor bovine serum in an incubator with 5% CO_2 at 37°C and were plated on polyacrylamide substrate overnight before TFM measurements.

Transfection of cells for the expression of GFP-Zyxin

GFP-zyxin plasmids were kindly provided by Dr. Juergen Wehland (Helmholtz Centre for Infection Research, Braunschweig, Germany). NIH 3T3 cells were cultured to reach 90% confluency on 60-mm tissue culture dishes and transfected with 2 μg of the plasmid using an Amaxa Nucleofector and Kit R (Lonza, Walkersville, MD), following the manufacturer's recommended protocol. Transfected cells were plated on gelatin-coated polyacrylamide substrates for TFM measurements.

Measurement of substrate displacements

To measure substrate displacements caused by traction forces, isolated NIH 3T3 cells were identified, and images of fluorescent beads were recorded with a Nikon Eclipse Ti microscope equipped with a dry 40×/N.A. 0.75 PlanFluor phase-contrast objective lens before and after removing the cell by adding 1 mL of 0.25% Trypsin-EDTA (Thermo Fisher Scientific, Waltham, MA) to the culture chamber. Images were captured with a Photometrics Prime 95B CMOS camera with 1200 × 1200 camera pixels (Tucson, AZ). At the magnification used, each camera pixel represented an area of 0.275 × 0.275 μm^2.

Bead images were cropped to a size of 728 × 728 camera pixels, with the cell positioned near the center surrounded by a margin such that bead displacements dissipated to near zero around the edge. Pairs of bead images before and after cell removal were analyzed with iterative particle image velocimetry to obtain the displacement field, using an open-access plugin for ImageJ (26). The particle image velocimetry program was configured to generate a displacement vector every 7 × 7 pixels. To obtain tightly packed vectors as input for DL-TFM, the sparse vector field was compacted by 7× along x and y directions to generate fields of 104 × 104 modeling pixels. Thus, each modeling pixel equaled 7 × 7 camera pixels covering an area of 1.925 × 1.925 μm^2. The 7× scaling factor was estimated based on a measured surface bead density of 18,900 ± 3050 beads over a field of 1200 × 1200 camera pixels or one bead per area of 8.7 × 8.7 pixels.

To measure and correct the residual alignment errors, a 9 × 9 median filter was applied to displacements over the area at least 35 pixels away from the edge of the cell. The average of filtered displacement vectors was taken as the residual alignment error and was subtracted from all the displacement vectors. Other defects of the displacement field often manifested as singular vectors that stood out against an otherwise slowly changing field. These defects were identified based on the ratio between the fields after convolving with a Laplace kernel and with a kernel that calculates the average of the immediate neighbors. Pixels with a ratio ≥3.5 were identified as defective and were replaced with the average displacement of surrounding pixels. This threshold was determined based on the maximal ratio that appeared in simulated defect-free displacement fields.

Statistical analysis

TFM was tested with 17 simulated cells of amoeboid or mesenchymal shapes, 16 simulated developing neurons, 10 simulated keratocytes, and 14 experimental NIH 3T3 cells. For simulated cells, the accuracy of TFM was assessed by comparing predicted stress field against the known ground truth field, using normalized RMSE defined as follows:

$$E = \sqrt{\left(\sum |V - V_{gt}|^2\right)} \Big/ \sqrt{\left(\sum |V_{gt}|^2\right)}, \qquad (1)$$

where V_{gt} is ground truth stress vectors, and V is the corresponding predicted stress vectors. The summation was carried out over all the modeling pixels within the border of a cell. Normalized RMSE was then averaged over testing cells within the group. Two-tailed *t*-test for paired samples was used for determining the statistical significance between two measurements. Because ground truth stress fields were not available for experimental measurements, predicted stress fields of experimental cells were assessed by comparing measured displacements against displacements calculated from the predicted stress field, using an equation similar to that shown in Eq. 1.

RESULTS

Construction and training of a modified U-Net for TFM

Vector fields of stress or displacement were represented as tensors of [x, y, v], where x and y were positions in a field of S × S modeling pixels (to be distinguished from camera pixels), and v ∈ {1, 2} represented the two orthogonal vector components. S ∈ {104, 160, 256} represented the size of three fields with the corresponding neural network. The original architecture of U-Net (22) was modified for processing the vector field of stresses and displacements by replacing 2D convolution operations with 3D convolution as required to create linear combinations of x and y components of stress vectors for generating the displacements in an elastic network (2). The resulting neural network architecture contained a total of 51 layers with three encoding levels for progressively decreasing the resolution and four decoding levels for restoring the resolution and three concatenation skip connections for combining the activations from corresponding encoding and decoding layers to preserve the spatial resolution ((22); Fig. S1, *downward vertical arrows* represent skip connections).

To overcome the challenge of obtaining large and precise data sets for training, we adapted a computational model for cell migration to generate 184 simulated amoeboid or mesenchymal cells (Fig. 1, *a–c*; (28)). To enhance radial symmetry of the neural network, each image was augmented by rotating seven times at 45° each. The training set also included 81 radial patterns of stress vectors spaced one pixel apart for the enhancement of resolution and radial symmetry (Fig. 1 *d*). Although the training process took hours to finish, the prediction of traction stress took 31–136 ms per field depending on the size (Table S1).

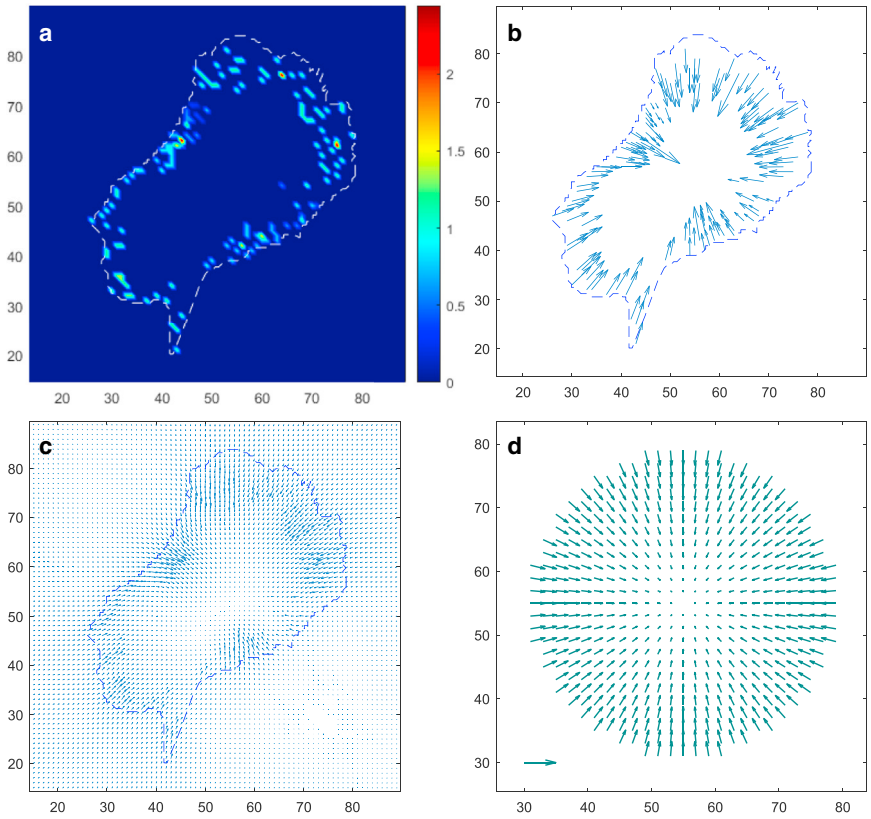

FIGURE 1 An example of stress and displacement fields for training the neural network. Heat map (*a*) and vector plot (*b*) show the traction stress field. Color scale is in the unit of Newtons/pixel^2. Displacements (*c*) are calculated assuming a Poisson ratio of 0.45 and Young's modulus of 10 N/pixel^2. Lengths of displacement vectors in pixels are enlarged five times (*c*). Segmented lines indicate cell border. To enhance radial symmetry and resolution of the neural network, the training set also includes radial patterns of stress vectors spaced one pixel apart (*d*). Bar, 5 N/pixel^2.

Testing DL-TFM and FTTC with simulated stress and displacement fields

To test if the trained neural network was able to generate correct predictions under conditions mimicking those for experiments, we placed simulated cells on a hypothetical substrate of 10,670 Pa as measured for polyacrylamide substrates of 8% acrylamide and 0.1% bis-acrylamide, in which they exerted a maximal stress of 6000 Pa representative of NIH 3T3 cells. Heat maps and vector plots of predicted and ground truth stress fields showed a strong qualitative resemblance (Fig. 2, *a, b, f, and g*). Quantitative assessments were based on normalized RMSE, defined as the RMSE between predicted and ground truth stress, divided by the root mean-squared magnitude of traction stress within the cell border. This approach allowed better comparison of cells with different magnitudes of traction stress while avoiding the dominance of error estimation by very small stresses that occupied most of the cell area.

The normalized RMSE averaged over 17 simulated testing cells was 10.8 ± 0.8% without adding noise. Imposing noise on displacement vectors at a magnitude comparable to that in experimental data increased the average normalized RMSE to 16.6 ± 0.68% without visible changes to predicted stress vectors (Figs. 2, *c and h*, and 3 *a*). The main effect of noise was an increase in small background vectors (Fig. 2 *h*).

As a comparison, we applied an open-access implementation of FTTC to the same data set (15). We first used simulated data to determine the optimal value of L2 regularization parameter λ. The method used was similar to the L-curve criterion but based on the difference between predicted and ground truth stress fields ((18); Fig. S2). The value of λ was determined as 1.6×10^{-10} for noise-free displacements and 3.5×10^{-10} after imposing noise.

With optimal λ, the FTTC algorithm tested also generated stress fields that appeared qualitatively similar to the ground truth and to the prediction of DL-TFM with a slightly higher background (Fig. 2, *d and i*). The average normalized RMSE for FTTC was 14.1 ± 0.22% for noise-free data and 21.5 ± 1.56% after imposing noise to displacement vectors (Fig. 3 *a*), both significantly above the corresponding normalized RMSE for DL-TFM. Interestingly, using predicted traction fields for reconstituting substrate displacements, we found that the FTTC algorithm tested showed a significantly smaller normalized RMSE than DL-TFM relative to ground truth displacements (Fig. 3 *a*, *rightmost bars*). The heat map of normalized errors for DL-TFM-predicted traction stresses showed that most prominent errors were located away from the cell perimeter and at the shoulder of stress peaks (Fig. 3 *d*), whereas errors from FTTC predictions were scattered more randomly across the cell (Fig. 3 *f*). In contrast, errors of reconstituted displacements showed similar patterns between DL-TFM and FTTC predictions (Fig. 3, *h–k*), with large errors concentrated near the cell center where stress vectors converge. The higher displacement errors for DL-TFM than FTTC came from the presence of low-level noise throughout the cell (Fig. 3, *h and i*), in addition to large errors near the cell center when noise was imposed on displacements (Fig. 3 *i*).

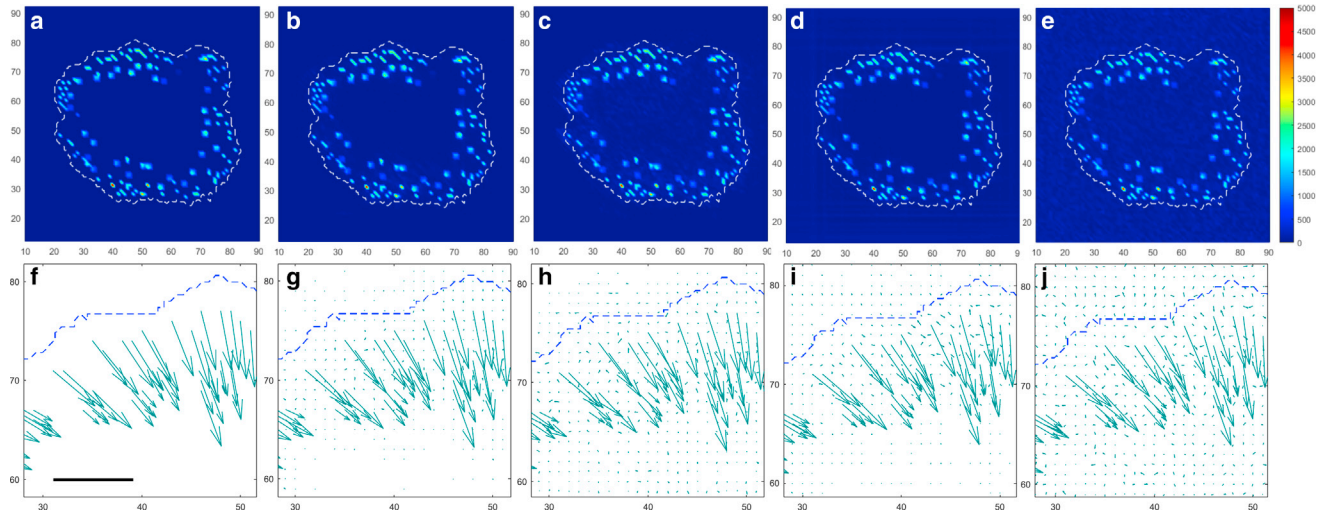

FIGURE 2 Traction stress fields of a simulated cell as predicted by DL-TFM (*b*, *c*, *g*, and *h*) or FTTC (*d*, *e*, *i*, and *j*). Magnitudes are rendered as heat maps to provide an overview (*a–e*). The fields are 104 × 104 pixels in size. Vector plots of the top left region of the cell are enlarged to provide a detailed view (*f–j*). DL-TFM, when applied to displacement fields without (*b* and *g*) or with (*c* and *h*) imposed noise, yields stress fields qualitatively indistinguishable from the ground truth (*a* and *f*). However, small noise vectors appear in the background (*g*), which become more pronounced upon the addition of noise to displacement vectors (*h*). Heat maps (*d* and *e*) and vector plots (*i* and *j*) of stress field generated by FTTC without (*d* and *i*) or with imposed noise (*e* and *j*) are also qualitatively similar to those of ground truth (*a* and *f*). FTTC is performed with a regularization parameter λ of 1.6×10^{-10} for noise-free data (*d* and *i*) and 3.5×10^{-10} after adding noise to displacements (*e* and *j*). Color scales for stress heap maps are in the unit of Pascals. Bar, 4000 Pa (*d*).

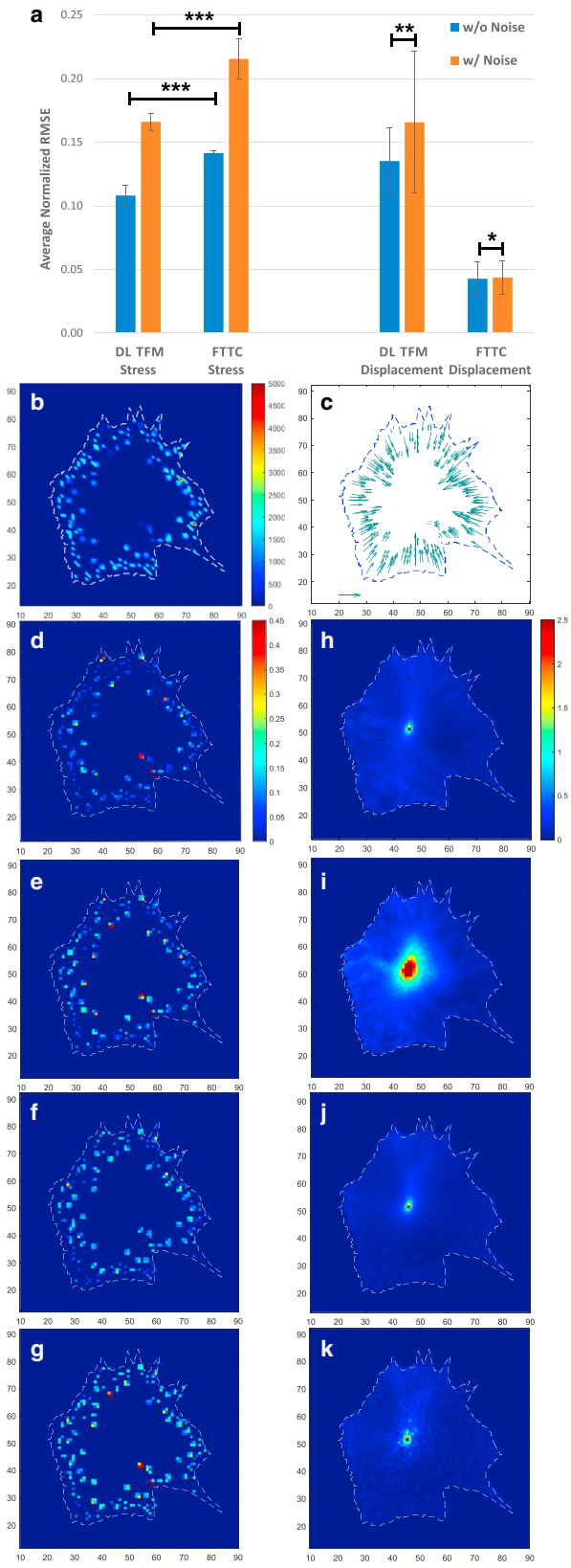

FIGURE 3 Normalized RMSE for DL-TFM and FTTC when applied to simulated cells with shapes of generic amoeboid or mesenchymal cells. Average normalized RMSE of traction fields predicted by DL-TFM or FTTC from 17 simulated cells are shown without (*blue bars*, *left side*) or with noise imposed on displacement vectors (*orange bars*, *left side*). L2 regularization of FTTC is performed with $\lambda = 1.6 \times 10^{-10}$ for noise-free data and 3.5×10^{-10} for noise-imposed data. Predicted displacements are then calculated from predicted traction fields without (*blue bars*, *right side*) or with imposed noise (*orange bars*, *right side*) and compared with the original displacements. As judged by the normalized RMSE of predicted traction fields relative to the ground truth (*left side*), DL-TFM shows a higher accuracy than FTTC. However, FTTC generates predicted displacements fields of smaller normalized RMSE than DL-TFM, regardless of the noise (*right side*). Error bars represent SDs. $*p > 0.05$, $**0.01 < p < 0.02$, $***p < 0.001$. All other adjacent pairs of normalized RMSE are significantly different with $p < 0.001$. The distribution of error is visualized by rendering the magnitude of error normalized with the magnitude of ground truth, for traction fields predicted with DL-TFM (*d* and *e*) or FTTC (*f* and *g*; color scale unitless) and for displacement fields reconstituted from predicted traction stress of DL-TFM (*h* and *i*) or FTTC (*j* and *k*; color scale unitless). Simulated displacements are used without (*d*, *f*, *h*, and *j*) or with (*e*, *g*, *i*, and *k*) imposed noise. Ground truth traction stresses are shown as a heat map (*b*; color scale in Pascals) or vector field (*c*) (bar, 4000 Pa). Traction stresses show a lower error when predicted with DL-TFM (*d* and *e*) than with FTTC (*f* and *g*). In contrast, displacements show a lower error when reconstituted with traction stresses predicted with FTTC (*j* and *k*) than with DL-TFM (*h* and *i*). Errors of reconstituted displacements are concentrated at the cell center, where traction vectors converge. DL-TFM shows a higher background error of displacements throughout the cell (*h* and *i*, *light blue background*).

Indicative of a high inherent resolution for both DL-TFM and FTTC, we noticed the preservation of pulses and closely spaced pairs of stress vectors in the predictions of simulated cells. Using a testing matrix of stress vectors in which adjacent stress vectors were separated by a single null pixel, we found that the pattern was well resolved for both DL-TFM and FTTC, with or without imposing noise to displacement vectors (Fig. 4, *b–e*). Compared with DL-TFM, FTTC showed more noise spikes on the background for data of high noise (Fig. 4, *c and e*). However, when the contrast of stress against the background was dropped from 100 to 20%, FTTC showed a better preservation of stress magnitude and resolution than DL-TFM (Fig. 4, *g–j*). Together, these results suggested that the limiting factor for TFM resolution was not the algorithm itself but conditions such as the density of beads on the substrate and the quality of images.

Scalability of DL-TFM

Scalability determines if a trained neural network is able to work under a range of imaging or mechanical conditions. To test the performance at different magnifications, we resized simulated cells and their displacement fields by a factor 0.8× or 1.25× using bilinear interpolation while maintaining a field size of 160 × 160 pixels (Fig. 5 *a*). Application of DL-TFM to resized displacement fields generated predicted traction fields that matched the resized ground truth with a normalized RMSE <20% without noise and <25% with noise (Fig. 5 *a*). In a second test, stress fields of 160 × 160 pixels were resampled to 104 × 104 or

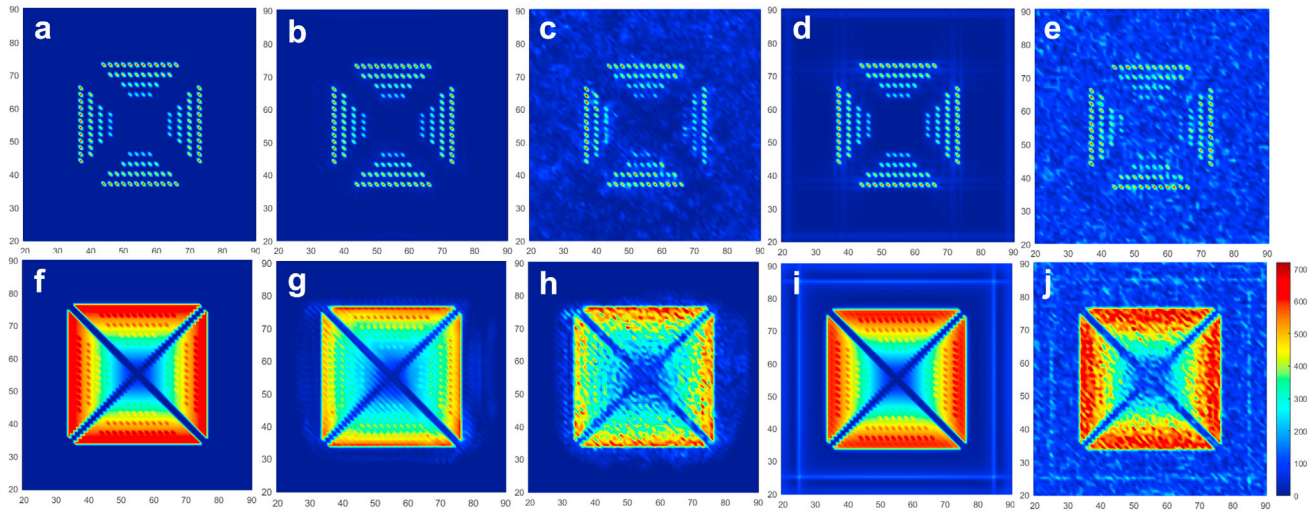

FIGURE 4 Resolution test for DL-TFM and FTTC at 100% (*a–e*) or 20% (*f–j*) contrast. The testing pattern consists of arrays of pulse stresses separated by one pixel between adjacent pulses. Heat maps of ground truth (*a* and *f*) and stress fields predicted by DL-TFM (*b*, *c*, *g*, and *h*) and FTTC (*d*, *e*, *i*, and *j*) indicate that even at a low contrast (*g* and *i*), both DL-TFM (*g*) and FTTC (*i*) are able to resolve at least part of the stress pulses when there is no noise. However, DL-TFM shows a decrease in stress magnitude (*g*). Adding noise to the displacements at low contrast abolished the resolution for both DL-TFM (*h*) and FTTC (*j*), whereas prominent noise artifacts appeared in the background predicted by FTTC (*e* and *j*). FTTC is performed with a regularization parameter $\lambda = 1.6 \times 10^{-10}$ for noise-free displacements and 3.5×10^{-10} for displacements with a noise of 7.65×10^{-3} modeling pixels. Color bars show the scale for stress in Pascals.

256×256 pixels while maintaining the size of the cell relative to the size of the field. The normalized RMSE of predicted traction fields was $\leq$20% without noise and <30% with noise, relative to resampled ground truth (Fig. 5 *c*). Although these tests were affected by the accuracy of interpolation, the results supported the notion that a trained neural network for DL-TFM is applicable to a range of cell sizes, imaging resolutions, or magnifications. We have also verified that the predictions were unaffected by the position of the cell relative to the center of the field (Fig. S3).

We then tested the scalability of DL-TFM under different mechanical conditions by first rescaling traction stresses of simulated cells for different maximal magnitudes between 1000 and 8000 Pa (Fig. 5 *d*). Normalized RMSE remained below 20% without imposed noise but increased progressively with decreasing maximal stress in the presence of noise as the noise became more prominent relative to the decreasing displacements. Similarly, the average normalized RMSE remained low for Young's moduli over a range of 2500–40,000 Pa without noise (Fig. 5 *b*) but increased with Young's modulus in the presence of noise, which became prominent relative to the small displacements on stiff substrates.

Finally, we tested the applicability of neural networks trained with simulated amoeboid and mesenchymal cells to other cell shapes, including the wedge shape of keratocytes and spiky shape of developing neurons, which were generated by adjusting parameters of the simulation model (28). Predicted stress fields remained qualitatively indistinguishable from the ground truth when rendered as either heat maps or vector plots (Fig. 6). The average normalized RMSE was similar to the values for simulated amoeboid and mesenchymal cells at 8.1 $\pm$ 0.76% for simulated keratocytes without noise and 14.2 $\pm$ 2.0% with added noise and 10.9 $\pm$ 2.2% for simulated neurons without noise and 19.2 $\pm$ 4.7% with added noise (Fig. 6 *i*), indicating that the accuracy was unaffected by the differences in cell shape and stress distribution. The FTTC algorithm tested also showed similar normalized RMSE for simulated mesenchymal cells (Fig. 3), keratocytes, and neurons (Fig. 6 *i*) under noise-free conditions. However, the error became prominent after imposing noise (Fig. 6 *i*). Although this was in part because of the lower traction stress and substrate displacements for simulated keratocytes and neurons than mesenchymal cells, the smaller error of DL-TFM for the same input data suggested that FTTC was more prone to noise than DL-TFM. Together, these tests suggested that DL-TFM is applicable to a wide range of cells and experimental conditions.

Application of DL-TFM to experimental data

We applied DL-TFM to experimental displacements obtained with NIH 3T3 cells on polyacrylamide substrates of 8% acrylamide and 0.1% bis-acrylamide (Young's modulus = 10,670 Pa). Displacement field was first corrected for residual misalignments of beads and then for defects that stood out as singular vectors against the surrounding by replacing them with the average of surrounding displacements (see Materials and methods).

Lacking ground truth, the accuracy of TFM was assessed based on normalized RMSE between measured displacements and displacements calculated from the predicted

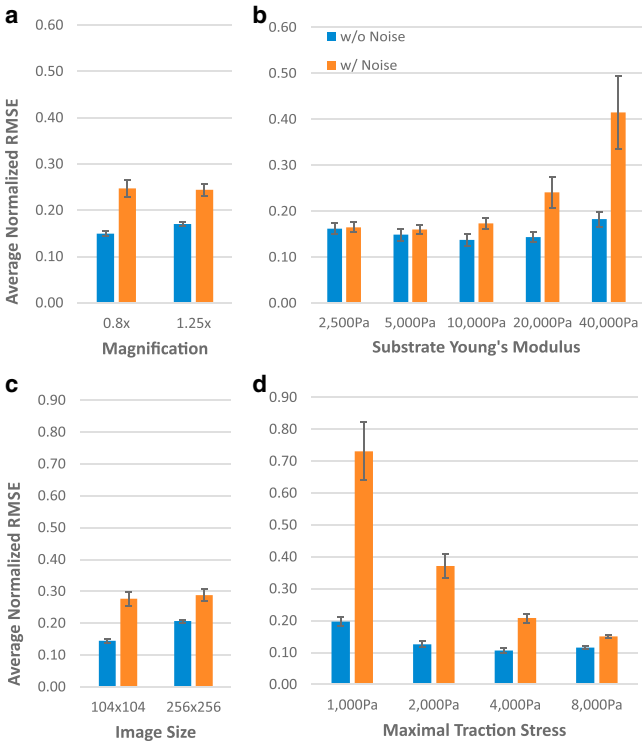

FIGURE 5 Scalability of DL-TFM for different spatial and mechanical conditions. Starting with a simulated cell in an area of 160 × 160 pixels, the size of the cell and its displacement field are resized by either 1.25× or 0.8× using bilinear interpolation without (*a*, *blue bars*) or with (*a*, *orange bars*) imposing noise. Application of DL-TFM to the resized displacement fields yields predicted stress fields that agree with resized stress field. Similarly, when an area of 160 × 160 pixels with a simulated cell is resampled by bilinear interpolation to 104 × 104 or 256 × 256 pixels (*c*), application of DL-TFM to resampled displacement fields yields predicted stress fields that agree with resampled stress fields without (*c*, *blue bars*) or with (*c*, *orange bars*) imposed noise. To test the scalability of mechanical conditions, simulated cells in an area of 160 × 160 pixels is placed on substrates with Young's moduli of 2500–40,000 Pa to generate displacements of different magnitudes (*b*). Application of DL-TFM to these displacement fields without noise generates stress fields that agree with the ground truth (*b*, *blue bars*). Imposing noise on the displacement field leads to unacceptable results for the very stiff substrate of 40,000 Pa (*b*, *tall orange bar*) because of its very small displacements. Other substrates yield acceptable results (*b*, the first four *orange bars*) based on a threshold of 40%. Finally, stress field, generated by a simulated cell on a substrate of 160 × 160 pixels with a Young's modulus of 10,670 Pa, is rescaled such that the maximal traction stress is 1000, 2000, 4000, or 8000 Pa (*d*). DL-TFM generates a predicted stress field that agrees with the corresponding rescaled stress field in the absence of noise (*d*, *blue bars*). Adding noise to the displacement field did not affect the validity of the predicted stress field based on a threshold of 40%, except for the weakest stress (*d*, first *orange bar*). Error bars represent SDs.

stress field. Heat maps and vector plots of stress fields generated by DL-TFM and FTTC showed qualitatively similar patterns (Fig. 7, *a and b*). The average normalized RMSE from 14 NIH 3T3 cells was 27.7 ± 7.2% for DL-TFM and 36.8 ± 12.4% for FTTC ($0.01 < p < 0.02$), using $\lambda = 1 \times 10^{-9}$ as the L2 regularization parameter for FTTC after testing with a range of λ values. Heat maps of normalized RMSE between predicted and measured displacement fields showed that the region of error was located typically in the area where backward-pointing stresses transitioned into forward-pointing stresses for both DL-TFM and FTTC (Fig. 7, *e–h*). Although these comparisons based on measured and predicted displacements suggested a similar accuracy between DL-TFM and the FTTC algorithm tested, the results with simulated cells pointed to the possibility that traction fields obtained with DL-TFM may have a high accuracy (Fig. 3).

We further compared the distribution of predicted traction stress and focal adhesions (Fig. 8). Although all the prominent focal adhesions, as seen in the image of GFP-zyxin (Fig. 8 *b*), were identified as sites of strong traction stress (Fig. 8 *c*), the relationship between the size of focal adhesion and magnitude of traction stress was known to be nonlinear (3,4). In addition, some sites of strong traction stress appeared to be at a short distance away from the cell border, which possibly reflected traction forces exerted by fine extensions such as filopodia or retraction fibers not easily visible in low-magnification images.

DISCUSSION

TFM is characterized mathematically as an ill-posed inverse problem (18), in which multiple solutions may meet the accuracy requirement while showing various degrees of artifacts. Although different methods have been developed for TFM, they followed related linear algebraic approaches, in conjunction with regularization to address the ill-posed nature (18). As a result, they face the common challenge of balancing between the accuracy of fit and the minimization of artifacts, with no definitive way to define the optimal regularization method or the regularization parameter λ (19).

Deep learning has emerged as an appealing general approach for solving ill-posed inverse problems (20). Although deep learning also applies regularization, it typically takes place only during the training stage (21) to prevent overfitting by balancing the goodness of fit for the training data against the complexity of weights in the neural network. As a result, deep learning avoids much of the compromises of regularization that happens at the prediction stage of conventional TFM. A second challenge for FTTC, and to a lesser extent for DL-TFM, is the presence of singular elements in Green's tensor. It affects FTTC directly during the prediction but DL-TFM only during the generation of training data, which supports the possibility that a well-trained network for DL-TFM can generate more accurate predictions than FTTC.

Successful applications of deep learning hinge on two major requirements, a network architecture compatible with the problem and the availability of a sizeable and reliable training data set. The deployment of U-Net, which was used widely for segmenting 2D images while preserving image quality, required important modifications to properly

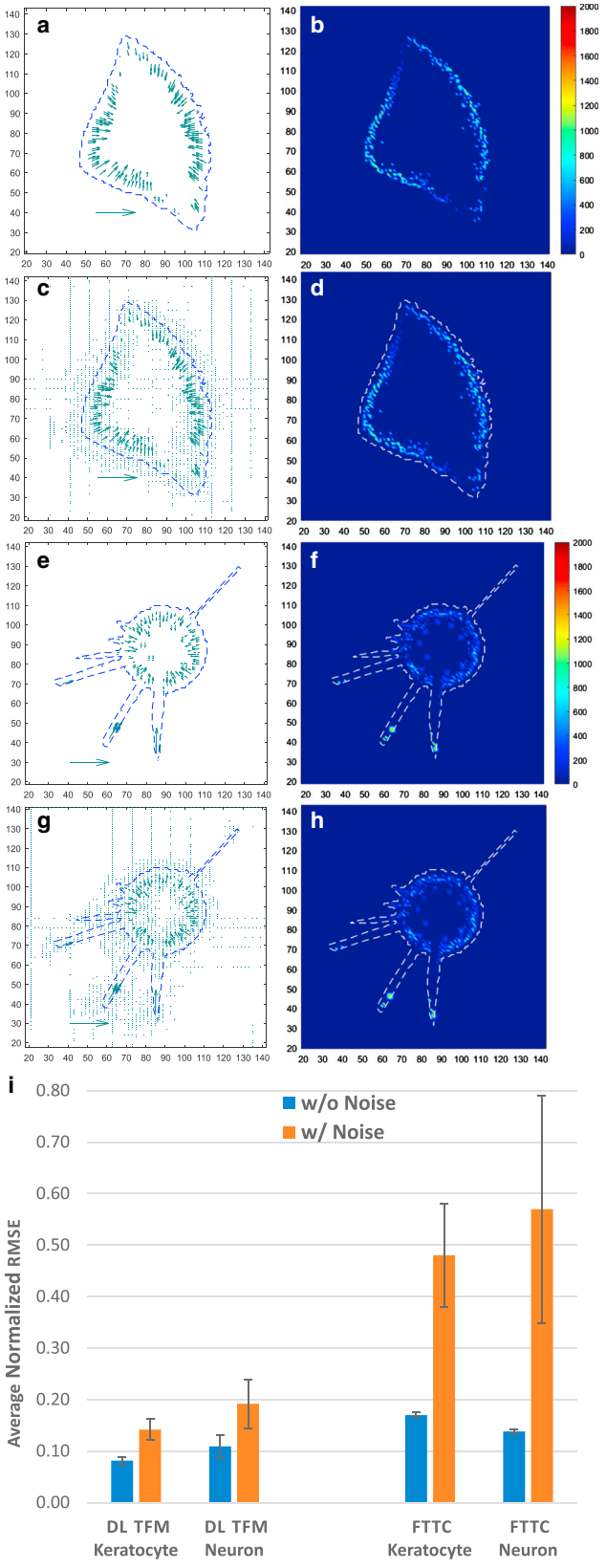

FIGURE 6 Applications of DL-TFM to simulated cells with shapes resembling keratocytes (*a–d*) or developing neurons (*e–h*). The same model that created simulated amoeboid or mesenchymal cells are able to generate simulated keratocytes and neurons after adjusting the parameters. Ground truth stress fields are shown as heat maps (*b* and *f*) and vector plots (*a* and *e*). Application of DL-TFM to displacement fields generates predicted stress fields (*c*, *d*, *g*, and *h*) that are qualitatively similar to the ground truth (*a*, *b*, *e*, and *f*). Quantitative comparison of normalized RMSE with 10 simulated keratocytes and 16 simulated neurons without (*i*, *blue bars*) or with imposed noise to displacements (*i*, *orange bars*) suggests that DL-TFM yields more accurate predictions than FTTC, particularly after imposing noise (*i*, *orange bars*). Error bars indicate SDs. All pairwise comparison between DL-TFM and FTTC are statistically significant ($p < 0.001$). Scale bar, 4000 Pa (*a*, *c*, *e*, and *g*).

reproduce the relationship between vector fields of stress and displacement. First, the vector nature of stress and displacement requires the use of two z-planes for representation. Second, linear elastic theory indicates that these planes cannot be processed independent of each other using 2D convolution, as for channels in a color image. Instead, 3D convolution is required to generate the required linear combinations that describe the relationship between stress and displacements.

Although the deployment of DL-TFM has been reported (24), the use of experimental data solved by conventional methods for training had two major drawbacks. Not only was it tedious to collect sufficient training data required for the performance of a neural network, the limitations of conventional TFM as addressed above also made it difficult to obtain precise training data. As a result, the performance of DL-TFM trained with data generated by conventional TFM is unlikely to surpass the performance of conventional TFM.

To obtain a large, precise training data set, we have adapted a simulation model for cell migration to generate unlimited pairs of stress and displacement fields that mimicked the pattern seen in amoeboid or mesenchymal cells (28). Although such simulation inevitably involved simplifications, the similarity of predictions generated by DL-TFM and by the conventional approach of FTTC argued for the validity and effectiveness of this approach.

Simulated data also allowed rigorous tests of DL-TFM under a wide range of conditions. We found that DL-TFM yielded smaller errors for stress prediction than the FTTC algorithms tested but cannot rule out the possibility that an improved implementation of FTTC or other conventional TFM methods may out-perform DL-TFM in accuracy (30,31). The accuracy of DL-TFM likely has to do with its optimization-driven approach that tunes tens of millions of network weights, guided by pattern recognition, to minimize the deviation between prediction and ground truth. Although deep learning lacked the analytical clarity of conventional TFM, its unique approach also circumvented the direct impact of singularity and regularization as seen in conventional approaches.

Interestingly, traction predictions with FTTC showed a sizeable error even under noise-free conditions, which required a nonzero regularization factor for mitigation.

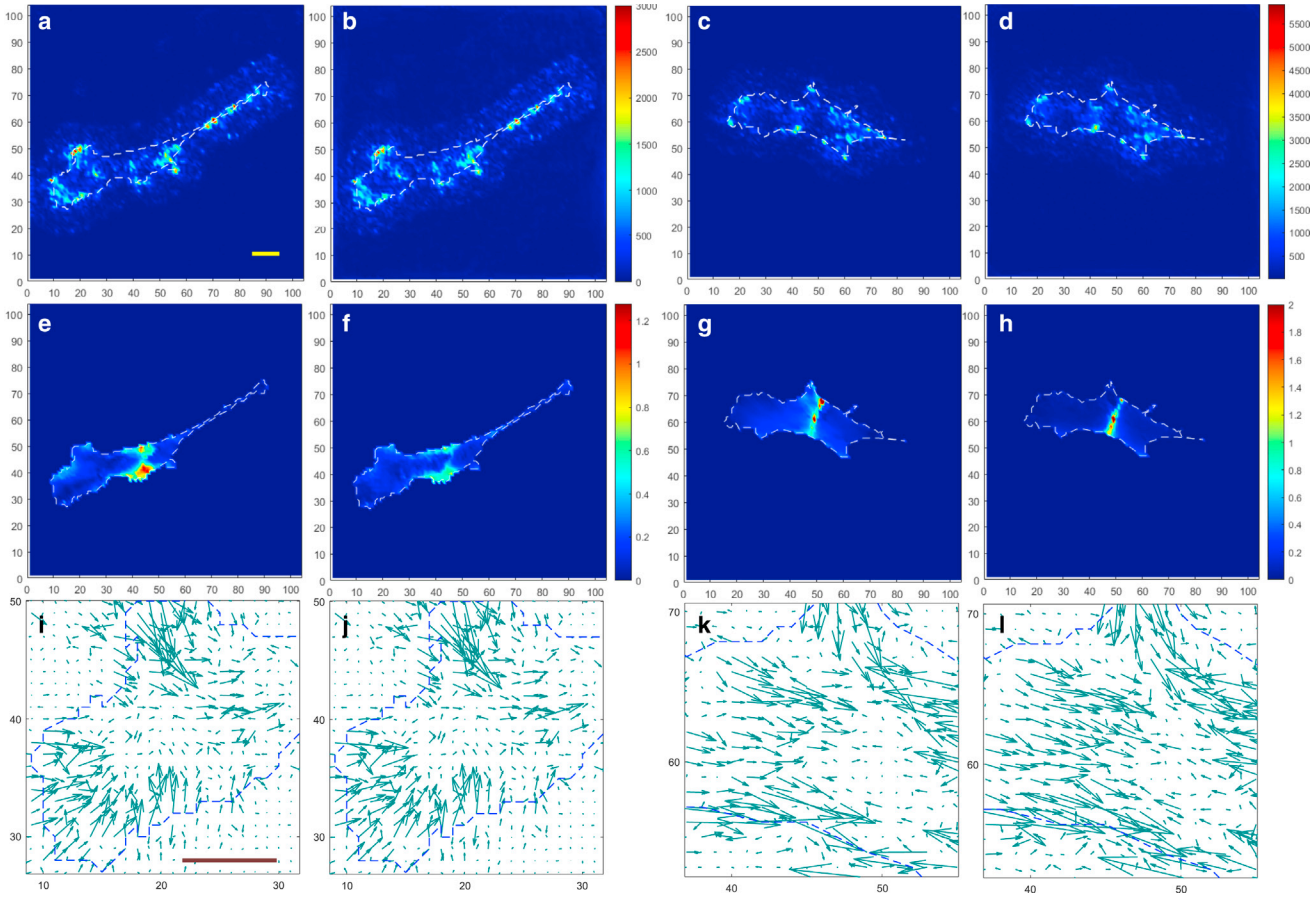

FIGURE 7 Traction stresses of two NIH 3T3 cells plated on polyacrylamide substrates of 10,670 Pa. The displacement field is analyzed with DL-TFM (*a*, *c*, *e*, *g*, *i*, and *k*) or FTTC with $\lambda = 1 \times 10^{-9}$ for L2 regularization (*b*, *d*, *f*, *h*, *j*, and *l*). Heat maps show stress fields predicted by DL-TFM (*a* and *c*) or FTTC (*b* and *d*) and the distribution of normalized RMSE between back-calculated and measured displacements for DL-TFM (*e* and *g*) and FTTC (*f* and *h*). Enlarged views of stress vectors show strong similarities near the front between DL-TFM (*i*) and FTTC (*j*). Differences between stress predictions generated by DL-TFM (*k*) and FTTC (*l*) are visible in the central region where large normalized RMSE are located, although their patterns remain qualitatively similar. The coordinates of vector plots serve as the guide for locating the corresponding regions in other images. Scale bar, 20 μm (*a*–*h*) or 4000 Pa (*i*–*l*).

The error likely originated from the approximation of the singular element at the center of Green's tensor. In addition, the calculation of traction stress in FTTC does not exactly represent the reciprocal of this approximation, which may leave some errors in the prediction. We were also intrigued by the observation that FTTC was more accurate than DL-TFM in recreating substrate displacements of simulated data, which may be related to the differential impact of the ill-posed nature of TFM on DL-TFM and FTTC. FTTC also showed a better preservation of stress magnitude and resolution when the stress field had a very low contrast. However, because traction forces are predominantly concentrated at focal adhesions, such low-contrast conditions arise only rarely in TFM.

Although DL-TFM understandably showed a slower speed than FTTC, the execution time remained well within 1 s for fields of up to 256 × 256 pixels using a commonly available computer. In addition, although DL-TFM required hours to train a neural network, well-trained networks proved applicable to a wide range of conditions, including cell size and shape, stress magnitude, image magnification, and substrate stiffnesses. Thus, DL-TFM may serve as a versatile alternative to conventional TFM. In addition, DL-TFM provides an independent approach whenever the prediction of conventional TFM is in doubt. Finally, the testing data set and strategy for the present study may prove valuable for the rigorous assessment of other TFM methods, whereas simulation may represent an appealing general approach for generating "big data" for machine learning, thereby avoiding the time-consuming collection of real-life data.

CONCLUSIONS

To address the challenge of TFM as an ill-posed inverse problem, we hereby introduced a new approach based on deep learning, which, when combined with large data sets generated by simulation, circumvented much of the challenges of conventional TFM. DL-TFM generated predictions qualitatively similar to those of the conventional

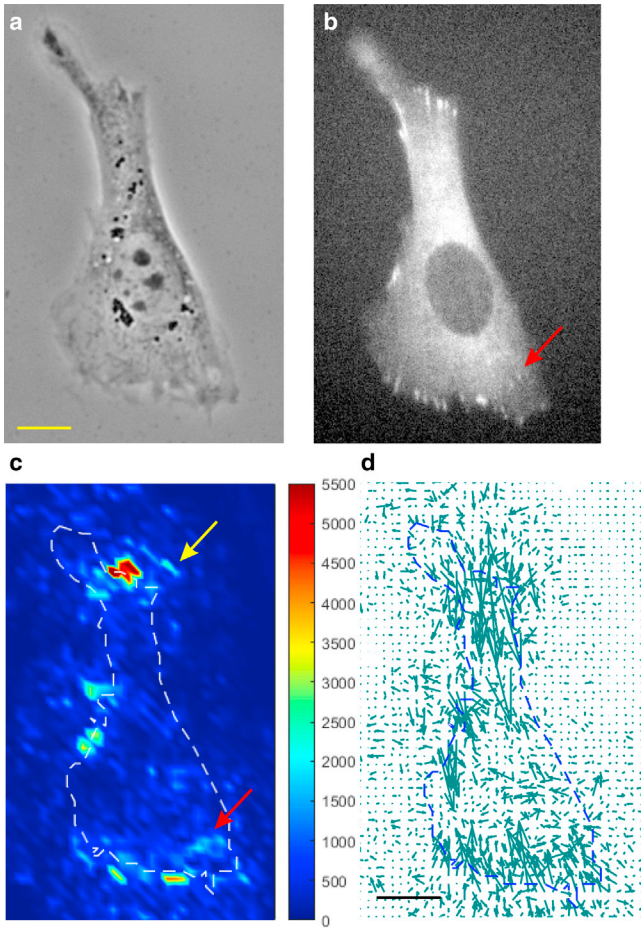

FIGURE 8 Relationship between traction stress predicted by DL-TFM (*c* and *d*) and focal adhesions imaged with transfected GFP-zyxin (*b*). Phase-contrast image (*a*) shows an NIH 3T3 fibroblast transfected with GFP-zyxin to label focal adhesions (*b*). Application of DL-TFM reveals a number of sites of strong stress, as shown in the heat map (*c*; color scale indicates stress in Pascals) and vector plot (*d*). Note that although all the prominent focal adhesions show prominent traction activities, the correlation between traction stress and focal adhesion size or intensity is not linear. Some traction stresses appear to colocalize with nascent focal adhesions (*b* and *c*; *red arrows*), whereas other strong stresses are at a distance from the cell edge (*c*; *yellow arrow*), possibly reflecting small adhesive extensions such as retraction fibers or filopodia. Scale bar, 20 μm (*a*) or 4000 Pa (*d*).

FTTC approach but at a higher accuracy than several implementations of FTTC tested using simulated testing data. DL-TFM also showed robust performances with regard to varying cell size, shape, magnification, substrate stiffness, stress magnitude, and stress pattern. We concluded that DL-TFM is a reliable and more accurate alternative to at least some of the conventional TFM methods for predicting cellular traction stress under a wide range of conditions.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2021.06.011.

AUTHOR CONTRIBUTIONS

Y.-l.W. performed all the computation, including the generation of simulated cells, the processing and analysis of experimental data, the construction and training of neural networks, and the performance of traction stress computation and evaluation. Y.-C.L. performed cell culture, prepared elastic substrates, performed transfections, and collected the experimental data.

ACKNOWLEDGMENTS

We thank Mr. Xinghan Li for constructing the preliminary version of the neural network, and Dr. Frederick Lanni for helpful discussions. This study was supported by National Institutes of Health grants R01 GM118998 and R35 GM136345.

REFERENCES

1. Sun, Z., S. S. Guo, and R. Fässler. 2016. Integrin-mediated mechanotransduction. *J. Cell Biol.* 215:445–456.
2. Dembo, M., and Y. L. Wang. 1999. Stresses at the cell-to-substrate interface during locomotion of fibroblasts. *Biophys. J.* 76:2307–2316.
3. Beningo, K. A., M. Dembo, …, Y. L. Wang. 2001. Nascent focal adhesions are responsible for the generation of strong propulsive forces in migrating fibroblasts. *J. Cell Biol.* 153:881–888.
4. Stricker, J., Y. Aratyn-Schaus, …, M. L. Gardel. 2011. Spatiotemporal constraints on the force-dependent growth of focal adhesions. *Biophys. J.* 100:2883–2893.
5. Lemmon, C. A., C. S. Chen, and L. H. Romer. 2009. Cell traction forces direct fibronectin matrix assembly. *Biophys. J.* 96:729–738.
6. Wong, S., W.-H. Guo, and Y. L. Wang. 2014. Fibroblasts probe substrate rigidity with filopodia extensions before occupying an area. *Proc. Natl. Acad. Sci. USA.* 111:17176–17181.
7. Rape, A. D., W.-H. Guo, and Y.-L. Wang. 2011. The regulation of traction force in relation to cell shape and focal adhesions. *Biomaterials.* 32:2043–2051.
8. Chang, S. S., A. D. Rape, …, Y. L. Wang. 2019. Migration regulates cellular mechanical states. *Mol. Biol. Cell.* 30:3104–3111.
9. Maruthamuthu, V., B. Sabass, …, M. L. Gardel. 2011. Cell-ECM traction force modulates endogenous tension at cell-cell contacts. *Proc. Natl. Acad. Sci. USA.* 108:4708–4713.
10. Baeyens, N., C. Bandyopadhyay, …, M. A. Schwartz. 2016. Endothelial fluid shear stress sensing in vascular health and disease. *J. Clin. Invest.* 126:821–828.
11. Mui, K. L., C. S. Chen, and R. K. Assoian. 2016. The mechanical regulation of integrin-cadherin crosstalk organizes cells, signaling and forces. *J. Cell Sci.* 129:1093–1100.
12. Lo, C. M., H. B. Wang, …, Y. L. Wang. 2000. Cell movement is guided by the rigidity of the substrate. *Biophys. J.* 79:144–152.
13. Wang, H. B., M. Dembo, and Y. L. Wang. 2000. Substrate flexibility regulates growth and apoptosis of normal but not transformed cells. *Am. J. Physiol. Cell Physiol.* 279:C1345–C1350.
14. Engler, A. J., S. Sen, …, D. E. Discher. 2006. Matrix elasticity directs stem cell lineage specification. *Cell.* 126:677–689.
15. Sabass, B., M. L. Gardel, …, U. S. Schwarz. 2008. High resolution traction force microscopy based on experimental and computational advances. *Biophys. J.* 94:207–220.
16. Schwarz, U. S., and J. R. D. Soiné. 2015. Traction force microscopy on soft elastic substrates: a guide to recent computational advances. *Biochim. Biophys. Acta.* 1853:3095–3104.
17. Colin-York, H., and M. Fritzsche. 2018. The future of traction force microscopy. *Curr. Opin. Biomed. Eng.* 5:1–5.

18. Schwarz, U. S., N. Q. Balaban, …, S. A. Safran. 2002. Calculation of forces at focal adhesions from elastic substrate data: the effect of localized force and the need for regularization. *Biophys. J.* 83:1380–1394.

19. Huang, Y., C. Schell, …, B. Sabass. 2019. Traction force microscopy with optimized regularization and automated Bayesian parameter selection for comparing cells. *Sci. Rep.* 9:539.

20. Adler, J., and O. Oktern. 2017. Solving ill-posed inverse problems using iterative deep neural networks. *Inverse Probl.* 33:124007.

21. Goodfellow, I., Y. Bengio, and A. Courville. 2016. Deep Learning. MIT Press, Cambridge, MA.

22. Ronneberger, O., P. Fischer, and T. Brox. 2015. U-Net: convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention, MICCAI 2015. Springer, pp. 234–241.

23. Alom, M. Z., C. Yakopcic, …, V. K. Asari. 2019. Recurrent residual U-Net for medical image segmentation. *J. Med. Imaging (Bellingham).* 6:014006.

24. Pielawski, N., J. Hu, …, C. Wählby. 2020. In silico prediction of cell traction forces. *In* IEEE 17th International Symposium on Biomedical Imaging (ISBI), pp. 877–881.

25. Butler, J. P., I. M. Tolić-Nørrelykke, …, J. J. Fredberg. 2002. Traction fields, moments, and strain energy that cells exert on their surroundings. *Am. J. Physiol. Cell Physiol.* 282:C595–C605.

26. Martiel, J.-L., A. Leal, …, M. Théry. 2015. Measurement of cell traction forces with ImageJ. *Methods Cell Biol.* 125:269–287.

27. Han, S. J., Y. Oak, …, G. Danuser. 2015. Traction microscopy to identify force modulation in subresolution adhesions. *Nat. Methods.* 12:653–656.

28. Satulovsky, J., R. Lui, and Y. L. Wang. 2008. Exploring the control circuit of cell migration by mathematical modeling. *Biophys. J.* 94:3671–3683.

29. Xiong, Y., C.-H. Huang, …, P. N. Devreotes. 2010. Cells navigate with a local-excitation, global-inhibition-biased excitable network. *Proc. Natl. Acad. Sci. USA.* 107:17079–17086.

30. Stricker, J., B. Sabass, …, M. L. Gardel. 2010. Optimization of traction force microscopy for micron-sized focal adhesions. *J. Phys. Condens. Matter.* 22:194104.

31. Suñé-Auñón, A., A. Jorge-Peñas, …, A. Muñoz-Barrutia. 2017. Full L_1-regularized traction force microscopy over whole cells. *BMC Bioinformatics.* 18:365.