Analyst

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: R. Guo, Z. Chen, X. Han, R. Sun, H. Lu, Y. Ma and L. Ma, *Analyst*, 2026, DOI: 10.1039/D6AN00232C.

View Article Online
DOI: 10.1039/D6AN00232C

ARTICLE

Data-Driven Optimization for Controllable Multi-Scale Aperture Fabrication of Nanopipettes

Runan Guo[a], Zhi Chen[a], Xue Han[a,e], Rongke Sun[a], Hui Lu[a,e], Yanqing Ma[a,b,c,d*] and Lei Ma[a,b,c*]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

This study addresses the critical challenge of controllable nanopipette fabrication by proposing a multi-parameter collaborative optimization framework that combines an artificial neural network (ANN) with a physical model. The ANN was utilized to establish a nonlinear mapping between the five strongly coupled fabrication parameters (Heat, Filament, Velocity, Delay, Pull) and nanopipette aperture size. Three complementary feature weight analysis methods—Random Forest (RF), SHAP values, and the Garson algorithm—were employed to rank the importance of these parameters, which consistently identified Heat as the most influential parameter, followed in that order by Pull, Delay, Filament, and Velocity. Concurrently, a physical model characterizing the temporal evolution of nanopipette aperture during the pulling process was derived from mechanical theory. Under the guidance of the theoretical model and parameter importance ranking, successful controllable fabrication of nanopipettes with target apertures spanning 50 nm to 1000 nm in 100 nm increments was achieved. This work thus transforms nanopipette fabrication from an empirical trial-and-error approach into a predictable, model-driven paradigm.

1 Introduction

Controllable fabrication at the nanoscale is a key challenge for various nanomaterial systems, including fluoride-based nanocrystals for three-dimensional stress imaging[1] and glass nanopipettes for single-cell analysis. Precise control of nanoscale features requires both advanced manufacturing tools and a thorough understanding of material performance under extreme conditions[2]. The practical application of nanopore sensors is often limited by poor fabrication consistency and significant device variability[3]. Nanopipettes serve as essential tools such as scanning electrochemical probes[4], instruments for microinjection[5], and vehicles for nanoscale drug delivery[6]. The size of their tip aperture directly determines measurement resolution and experimental applicability.

Different applications have specific and varied requirements for the aperture size of nanopipettes. Generally, nanopipettes with small apertures (50–300 nm) have high spatial resolution. They are mainly used for high-sensitivity single-molecule/single-cell biosensing[7, 8], such as current rectification-based detection of proteins[9] (50 nm), pH[10] (90 nm), and ATP[7] (220 nm). They also serve for precise operations, including DNA delivery[11] (100 nm) and intranuclear sampling[12] (300 nm).

On the other hand, larger-aperture nanopipettes (400–1000 nm) have higher flux or loading capacity[5]. They are mainly used for electrochemical imaging, such as SECM/SICM probes (440 nm). They are also applied to visualize material interface processes, such as bubble nucleation[13] (500 nm) and electrochemical reaction kinetics[14] (1000 nm). Additionally, they enable precise manipulation of microliter-scale liquids, such as cell injection[15] (600 nm). Therefore, different applications have clear and distinct aperture requirements for nanopipettes. Manufacturing nanopipettes with stable, multi-level apertures is challenging. Current methods for precisely controlling pulling parameters to achieve desired aperture sizes rely heavily on trial and error. This makes it hard to systematically produce nanopipettes of specific sizes as designed.

To obtain high-quality nanopipettes with controllable, and multi-scale dimensions, researchers have investigated the influence of pulling parameters on pipette shape. Traditional optimization often employs the single-factor method[16, 17], where one parameter is varied while others are held constant. Although this method is simple, it cannot explain the complex interactions between parameters. Furthermore, it is difficult to systematically optimize parameters for specific design goals. As a result, experimental processes rely heavily on experience-driven trial and error. Although some studies have explored multifactorial strategies like orthogonal arrays and Response Surface Methodology (RSM) to investigate parameter relationships and optimize microfabrication parameters, microfabrication limitations persist. Zhang et al[18] used orthogonal experiments to interpret two-factor interactions but

[a] Tianjin International Center for Nanoparticles and Nanosystems, Tianjin University, Tianjin University, 300072, P. R. China
[b] Tianjin Key Laboratory for Low-dimensional Electronic Materials and Advanced Instrumentation, Tianjin University, 300072, P.R. China.
[c] Haihe Laboratory for Low-dimensional Electronic Materials, Add 1 to No. 57, Wujiayao Street, Hexi District, Tianjin, 300074, P.R. China.
[d] School of Precision Instrument and Opto-electronics Engineering, Tianjin University, Tianjin 300072, P.R. China
[e] School of Chemistry and Chemical Engineering, Shihezi University, 832003, P.R. China

View Article Online
DOI: 10.1039/D6AN00232C

could only produce micropipettes with tip diameters of 10 µm and diameter variations less than 11.5 µm over 100 µm length. Wang et al[19] applied RSM to model the Length of the Tapered Vessel Segment (LTVS), revealing relationships between Heat, Velocity, Pull, and LTVS. However, the linear regression or polynomial fitting used in RSM has limited ability to describe nonlinear dynamic processes. By contrast, artificial neural networks (ANNs) have recently been applied to optimize thermomechanical properties in additive manufacturing, demonstrating strong predictive power for nonlinear process–property relationships[20]. Beyond the specific case of nanopipettes, the broader field of micro/nanofluidic devices has witnessed increasing adoption of artificial intelligence (AI) to address challenges in material processing, nanoscale control, and device fabrication[21]. For instance, in fused deposition modelling of carbon fiber-reinforced nylon composites, ANN model was used to predict tensile strength and surface roughness based on five printing parameters, achieving high accuracy ($R^2 > 0.99$) and outperforming other machine learning algorithms[22].

The key challenges for the controllable fabrication of multi-scale nanopipettes include two aspects. One is the unclear relative influence of each parameter on the target aperture and their complex coupling. The other is the lack of systematic studies on controllable fabrication across the full range of desired nanopipette sizes. While previous studies have explored single-factor analysis, orthogonal arrays, and response surface methodology, a systematic framework that combines data-driven modelling with physical insight to enable predictable, multi-scale nanopipette fabrication has not been demonstrated.

In this work, an ANN-based framework is proposed for nanopipette size prediction and parameter optimization. It first constructs a nonlinear mapping between five pulling parameters (Heat, Filament, Velocity, Delay, Pull) and tip aperture, overcoming traditional dimensionality constraints. Three complementary methods—Random Forest (RF), SHAP values, and the Garson algorithm—are utilized to quantify feature importance, clarifying clear optimization priorities. Concurrently, a mechanics-based physical model reveals the underlying aperture formation mechanism. Guided by the cross-validated parameter weights, 11 distinct nanopipette diameters (50–1000 nm in 100 nm increments) are successfully fabricated. These methodological steps (ANN mapping, multi-method importance analysis, physics-informed modelling) and the multi-scale controllable fabrication capability constitute the core contributions of this work.

2 Experimental

2.1 Materials

Borosilicate glass capillaries (Sutter Instrument Company, Model BF100-58-10) were used. The nominal inner diameter is 0.58 mm. The outer diameter is 1.0 mm. The length is 100 mm.

2.2 Nanopipette Fabrication

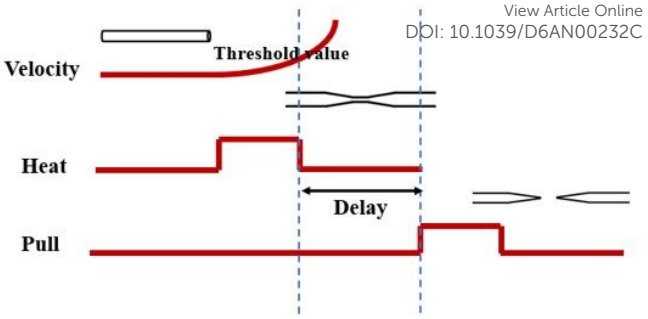

Figure 1 Relationship diagram of puller process parameters[23].

All nanopipettes were fabricated using the Sutter P-2000 laser puller (Sutter Instrument Company). This instrument employs a CO_2 laser to melt the glass capillary, which is clamped at both ends to puller solenoids. A scanning mirror directs the laser beam onto the capillary's back; light penetrating the capillary is refocused by a concave mirror behind it onto the front side, providing uniform circumferential heating. When the glass softens and reaches a preset velocity threshold, the solenoids activate, stretching and separating the capillary into two pipettes[23,24]. The relationship between these process parameters is shown in Figure 1.

The P-2000 puller has five programmable parameters. Heat values range from 0 to 999 and determine the laser power output of the instrument. A Heat setting below 100 results in 0 W output. Starting at 100, the power increases linearly by 1 W for every increase of 90 heat units. The maximum power of 10 W is delivered at the maximum Heat setting of 999. Filament value is between 0 and 5. It specifies the scanning pattern of the laser beam along the length of the glass capillary. Scan lengths are pre-programmed, ranging from 1 mm to 8 mm. Velocity value is between 0 and 255.

The Velocity parameter corresponds to a threshold voltage derived from an 8-bit DAC. Each unit of this parameter represents approximately 4 mV. Each unit of this parameter corresponds to a stretching velocity of approximately 0.025 mm/s. When the amplified velocity signal exceeds the threshold voltage corresponding to the Velocity parameter, it triggers the pull mechanism. Pull value is between 0 and 255. The Pull parameter controls solenoid-driven hard pull force, with current proportional to the setting. A setting of 255 corresponds to approximately 2.5 A. Since the electromagnetic force depends nonlinearly on both the current and the core position, the actual force output must be determined empirically. For example, at a setting of 157, the measured force is approximately 8.9 N, while at setting 221, it reaches approximately 22.24 N. Delay value is between 0 and 255. Delay (ms) regulates the time between laser shutdown and hard pull activation. A setting of 128 represents zero delay. Settings lower than 128 trigger the pull early (before the programmed heating duration ends), while settings higher than 128 increase the delay time (t_{delay}) before the pull starts[24].

The aperture morphology of all fabricated nanopipette samples was characterized by scanning electron microscopy (SEM, Hitachi, SU3500).

View Article Online
DOI: 10.1039/D5AN00372G

2.3 ANN and Random Forest Models

ANN Model

The ANN architecture consists of an input layer, three hidden layers (30, 20, 10 neurons) with Dropout regularization, and an output layer. The input layer contains five parameters, and the output layer predicts the aperture diameter[25, 26].

Hidden layer neurons utilize the Sigmoid activation function, which is mathematically defined by the equation:

$$\sigma(x) = \frac{1}{(1+e^{-x})} (1)$$

Sigmoid introduces essential non-linearity[27]. Input features were standardized using Z-score normalization during training. The dataset was randomly divided into training (80%) and validation (20%) to ensure objective model evaluation. The Adam optimizer (initial learning rate 0.001) was used for parameter updates. Mini-batch training was implemented to enhance gradient stability. Training proceeded for a maximum of 5000 epochs. Three Dropout layers (rate = 0.3) were inserted between hidden layers to randomly deactivate neurons during training. Performance metrics (Mean Squared Error, MSE) were calculated on the independent validation set every 10 epochs to monitor overfitting risk. Training aimed to minimize MSE[28]:

$$MSE = \frac{1}{N}\sum_{i=1}^{l}\left(y_i^p - y_i^t\right)^2 (2)$$

where n is the number of samples, y_i^p is the value predicted by the ANN for sample i, and y_i^t is the actual measured value for sample i. The model's performance was ultimately quantified through a quantitative evaluation of its predictive capability, using the MSE on the validation set as the loss function.

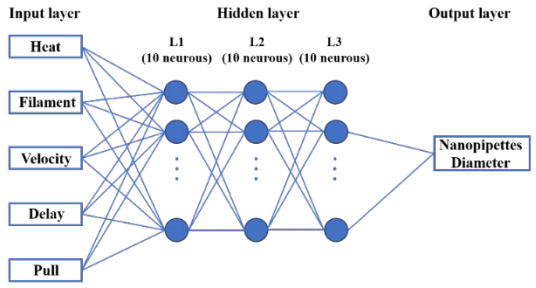

Figure 2 Structure of the Artificial Neural Network model.

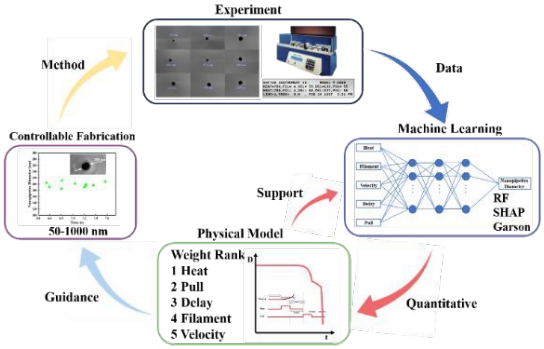

Figure 3 Flowchart illustrating the experimental methodology and optimization process.

Random Forest Model

Random Forest (RF) is an ensemble learning method. It builds multiple decision trees, which are based on random feature subsets and bootstrap sampling. Then it aggregates their prediction results. Its generalization error converges to a constant value, which is bounded by the strength of individual trees and the inter-tree correlation. RF is robust to high-dimensional data and noisy scenarios and provides variable importance measures[29]. To compare predictive performance and evaluate feature importance, an RF regression model was also constructed for predicting nanopipette aperture. The ensemble contained 1000 decision trees. The model employs a dual randomization to enhance robustness and diversity, comprising row randomization via bootstrap sampling and column randomization in the feature selection step. In row randomization, each decision tree is trained on a unique subset obtained through bootstrap sampling with replacement from the original training set. Concurrently, column randomization restricts feature candidates at each node splitting step to a randomly selected subset of size $k^{1/2}$ (k equals the total feature count of 5 fabrication parameters) when determining optimal splits. The model utilized Out-Of-Bag (OOB) error for evaluation. For each tree, samples not included in its bootstrap sample (OOB samples) form an independent validation set[29, 30]. The final model prediction error is estimated by calculating the MSE over all samples using only predictions from trees for which the sample was OOB:

$$OOB_{Error} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i^{OOB}\right)^2 (3)$$

Where n is the sample count, y_i is the actual aperture measurement of sample i, and y_i^{OOB} is the prediction for sample i averaged only over trees that did not have sample i in their bootstrap training set. This method efficiently utilizes all sample information for evaluation without requiring an explicit hold-out validation set, making it particularly suitable for our medium to small-scale dataset.

The overall workflow is summarized in Figure 3.

3 Results

3.1 Single-Factor Analysis

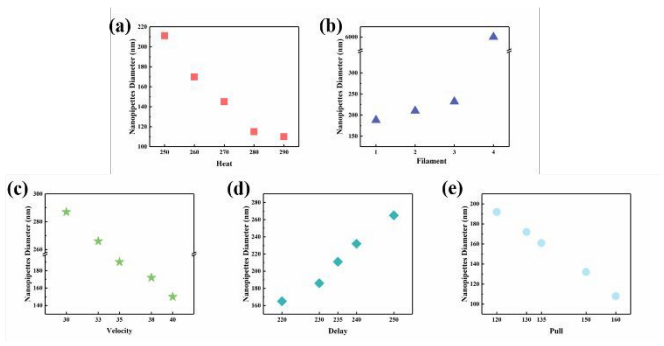

Figure 4 Influence of individual fabrication parameters on nanopipette aperture diameter (a) Heat, (b) Filament, (c) Velocity, (d) Delay and (e) Pull.

View Article Online
DOI: 10.1039/D6AN00373C

The influence of the five fabrication parameters on aperture was initially explored using the single-factor method. Results showed that increasing parameter values led to decreased nanopipette aperture for Heat (Figure 4a), Velocity (Figure 4c), and Pull (Figure 4e). Conversely, increasing Filament (Figure 4b) and Delay (Figure 4d) resulted in increased aperture. This confirms that Heat, Filament, Velocity, Delay, and Pull do not act in isolation but are intricately intertwined, collectively influencing the final aperture size.

3.2 ANN-Based Aperture Prediction and Feature Importance

Model training utilized a dataset comprising 135 sets of fabrication parameter combinations and their corresponding measured aperture values (Table S1). The frequency distributions of these 135 parameter sets are presented in Figure S2. As shown in Figure 5(a), during the training process, the loss values (Loss) of the training set and validation set exhibit a favorable convergence trend with the increase in iterations. For comparison, the loss curves of the ANN without Dropout are provided in Figure S3(a), which shows noticeable overfitting with a larger gap between training and validation loss. By contrast, the Dropout-regularized model (Figure 5a) yields more stable loss and better generalization. Figure 5(b) presents the comparison curve of pore diameters between the training set and test set, with a correlation coefficient R^2 of 0.94, where the correlation coefficient for the validation set is 0.98 (inset of Figure 5(b)). Figure 5(c) provides an intuitive comparison between the validation set predictions and the experimental values, whose trends are highly consistent, directly verifying the reliability of the model's predictions on new data. To ensure the robustness of feature importance analysis, three complementary methods—Random Forest (RF), SHAP values, and the Garson algorithm—were adopted for comparative analysis. Meanwhile, the out-of-bag (OOB) error was utilized to assess model performance during the training process.

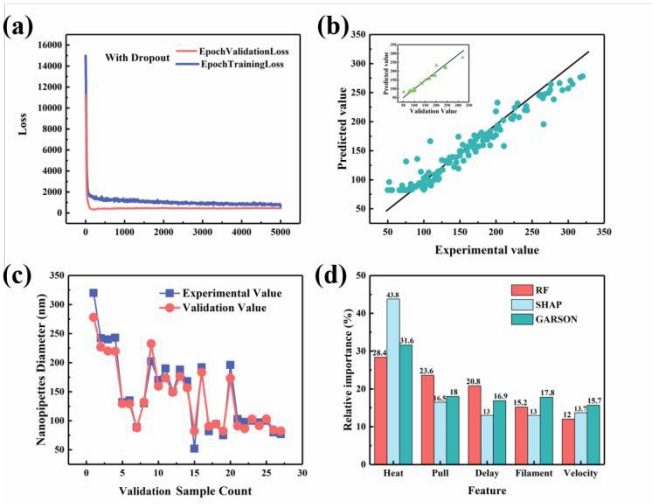

Figure 5 (a) Convergence curves of training and validation loss for the Dropout-regularized ANN over 5000 iterations;(b) Comparison curves of pore diameters between the training set and test set, with the inset showing the validation set;(c) Comparison curves between model-predicted validation values and experimental values;(d) Relative importance of input features with respect to the output nanopipette diameter.

As shown in Figure S3(b), the OOB mean squared error (OOB-MSE) decreased rapidly with the number of grown trees, converging to a stable value of approximately 1260 when the number of trees exceeded 200. The final RF model, trained with 1000 trees, achieved an OOB coefficient of determination (R^2) of 0.73, indicating a good fit between the predicted and measured nanopipette aperture sizes.

As summarized in Figure 5(d), all three methods consistently identify Heat, Pull, and Delay as the most critical parameters influencing aperture prediction. Specifically, Heat ranks first across all methods (RF: 28.4%, SHAP: 43.8%, Garson: 31.6%), followed by Pull (RF: 23.6%, SHAP: 16.5%, Garson: 18.0%) and Delay (RF: 20.8%, SHAP: 13.0%, Garson: 16.9%). Filament and Velocity show relatively lower importance.

3.3 Physical Model of Aperture Formation

Model Formulation

To understand the physical mechanisms governing nanopipette fabrication and to guide process optimization, a physical model was developed to describe the evolution of glass capillary tip aperture with pulling time. The model is based on the core assumption that softened glass behaves as an incompressible Newtonian fluid. Due to the rapid heating rate of the laser, the narrow temperature range for elastic deformation is neglected, and the model focuses only on plastic deformation occurring in the viscous flow state.

The core concept of the model relies on volume conservation, which involves approximating the deformation of the glass within the effective laser heating zone (L_{eff}) as a transformation from an initial cylinder to a final conical frustum (Figure 6). Let A_1 be the initial cross-sectional area of the heated zone, A_2 be the tip cross-sectional area at the end of deformation, and ΔL be the total elongation of the heated zone in the stretching direction.

The volume conservation equation is:

$$A_1 L_{eff} = \frac{1}{3}\left(L_{eff} + \Delta L\right)\left(A_1 + \sqrt{A_1 A_2} + A_2\right) \quad (4)$$

Defining $K = A_2/A_1$ (area scaling ratio) and solving for K yields:

$$K = \frac{A_2}{A_1} = \frac{1}{4}\left(\sqrt{\frac{12L_{eff}}{L_{eff} + \Delta L} - 3} - 1\right)^2 \quad (5)$$

The corresponding aperture diameter ratio is:

$$D_2 = \frac{D_1}{2}\left(\sqrt{\frac{12L_{eff}}{L_{eff} + \Delta L} - 3} - 1\right) \quad (6)$$

Where D_1 and D_2 are the aperture diameters at the start and end of deformation, respectively.

Glass viscosity η follows the Vogel-Fulcher-Tammann (VFT) equation with temperature:

$$lg\eta = A + \frac{B}{T - T_0} \quad (7)$$

where A, B, T_0 are material constants, and T is absolute temperature.

View Article Online
DOI: 10.1039/D5AN00272C

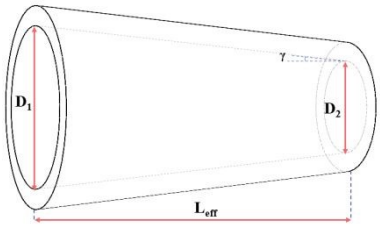

Figure 6 Schematic diagram of pipette structure during stretching. D_1 and D_2 represent aperture diameters at the start and end of deformation, respectively. Leff is the effective laser heating zone length.

When the capillary temperature exceeds the glass transition temperature T_g, entering the viscous flow stage, its strain rate ($\dot{\varepsilon}$) relates to applied stress (σ) via the Newtonian fluid constitutive equation:

$$\frac{d}{dt}\left(\frac{L(t)}{L_0}\right) = \frac{v(t)}{L(t)} = \dot{\varepsilon} = \frac{\sigma(t)}{3\eta(T)} \quad (8)$$

Here L(t) is a characteristic length (heated zone length or aperture-related length) at time t, L_0 is its initial length, v(t) is the stretching velocity. The tensile stress $\sigma(t)$ is determined by the pulling force F and the current cross-sectional area A(t):

$$\sigma(t) = \frac{F}{A(t)} \quad (9)$$

Combining Eq. (8) and (9), the relationship between stretching velocity, length, area, and viscosity is obtained:

$$v(t) = \frac{FL(t)}{3\eta(T)A(t)} \quad (10)$$

This leads to the differential equation for length change:

$$\frac{dL}{dt} = \frac{FL(t)}{3\eta(T)A(t)} \quad (11)$$

Solving this differential equation gives the expression for length evolution over time:

$$L(t) = e^{\int \frac{F}{3\eta(T)A(t)}dt - lnL_{initial}} \quad (12)$$

During the Heating Phase, $L_{initial}$ is the initial heated zone length, integration runs from start time to heating end time t. In the Cooling Phase, $L_{initial}$ is the length at the end of heating, integration runs from cooling start time to end time t. Substituting $\Delta L(t) = L(t) - L_{initial}$ from Eq. (12) into Eq. (6) yields the core model equation for aperture evolution over time:

$$D_2 = \frac{D_1}{2}\left(\sqrt{\frac{12L_{eff}}{L_{eff} + L(t)} - 3} - 1\right) \quad (13)$$

D_1 is the aperture at the start of the current phase (heating or cooling), and D_2 is the aperture at time t. This equation quantitatively characterizes the influence of key process parameters on aperture formation. Heat (Laser Power P) affects the heating rate, dT/dt, which is proportional to P and thus to the Heat value. Higher Heat means faster heating, which causes viscosity $\eta(T)$ to drop more rapidly and to a lower value. This leads to a larger value of the integral $\int[F/(3\eta A)]dt$ being achieved in a shorter time, which results in a smaller D_2. Pull (F) directly affects the numerator F in the integral of Eq. (11) during the cooling phase. A higher Pull (F_2) increases the integral value, thereby leading to a smaller D_2. Delay (Time) affects the starting temperature T_{start} of the cooling phase. The Delay value determines the time interval (t_{delay}) between laser shutdown and pull start, thereby influencing the degree of natural cooling during this period. Smaller Delay (t_{delay}) means higher T_{start} and lower initial viscosity $\eta(T_{start})$ in the cooling phase, which causes the integral term to grow faster initially, ultimately resulting in a smaller D_2. Delay-phase deformation remains negligible and is modeled as uniformly accelerated stretching. This approximation is valid because the force (F_1) is relatively small, and the delay time(t_{delay}) is very short.

Model Validation

To validate the physical model, experimental parameters were input to compare its output with expected trends. While precise quantitative prediction is challenging due to complex factors like surface tension, laser fluctuations, and environmental disturbances, the focus is on whether the predicted trends align with the actual physical process. Take the parameters in Table 1 as an example. The outer diameter (D_0) is 1 mm. The inner diameter (d_0) is 0.58 mm. The initial area (A_1) is 5.21×10^{-7} m^2. The initial heated zone length (L_{eff}) 3 mm. The heating phase force (F_1) is 1.4 N. The cooling phase force (F_2) is 5.6 N. The temperature parameters initial (T) is 25°C. The heating rate (dT/dt) is approximately 190°C/s. The laser power(P) is approximately 2.83 W. The spot area(A) is 9 mm^2. The density(ρ) is 2230 kg/m^3. The specific heat(c) is 800 J/(kg·K). The heating depth(d) is 0.1 mm. The cooling rate(dT/dt) is -736 °C/s. Total heat loss (Q) comprises convective and radiative components. The simulation yielded the nanopipette aperture evolution trend over time shown in Figure 7. It should be noted that the initial 2.7 seconds before the temperature reaches the glass transition point T_g are omitted.

Table 1 Example Simulation Parameters

Heat	Filament	Velocity	Delay	Pull	Time(s)
255	3	35	235	135	6.93

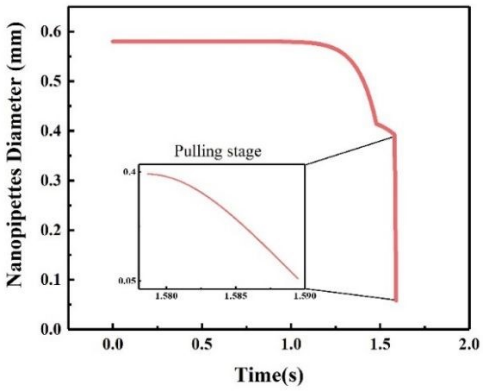

Figure 7 Simulated evolution curve of nanopipette aperture diameter over time (Time zero corresponds to the instant temperature reaches T_g). (Inset) Detail of aperture diameter kinetics during force application in the Pull phase.

View Article Online
DOI: 10.1039/D6AN00232C

3.4 Weight-Guided Optimization and Controllable Fabrication (50-1000 nm)

To verify the guiding value of the theoretical models and establish a standardized fabrication protocol, a verification study was conducted for 200 nm aperture fabrication via a feature weight-guided hierarchical optimization strategy. This aperture size was selected as a representative optimization example because, as a medium aperture spanning the full fabrication scale (50–1000 nm), its optimization methodology is generalizable to all target sizes and thus establishes a replicable paradigm for subsequent full-scale controllable fabrication. For this study, a baseline parameter set that yielded an aperture close to the 200 nm target (Table 2) was first selected from the initial experimental database.

Given the dominant feature importance of Heat (Table 3), it was optimized first. Ten nanopipettes were fabricated for each parameter set, with SEM characterization shown in Figures S4–S6. Statistical analysis of Figure 8 indicates Group 2 (Heat of 252) achieved optimal stability. This group exhibited the smallest interquartile range (IQR of 21 nm), a median of 200.5 nm (closest to the 200 nm target), a mean of 200.9 nm, and an average deviation of less than 3.5% from the target. Furthermore, no outliers were observed, and the boxplot exhibited a box height to whisker length ratio of 2.5.

Table 2 Initial Fabrication Parameters for 200 nm Nanopipette

Heat	Filament	Velocity	Delay	Pull
250	3	35	235	135

Table 3 Parameters for Optimizing Heat Value

Group	Heat	Filament	Velocity	Delay	Pull
1	250	3	35	235	135
2	252	3	35	235	135
3	255	3	35	235	135

Table 4 Delay/Pull Parameter Optimization Table

Group	Heat	Filament	Velocity	Delay	Pull
1	252	3	35	235	133
2	252	3	35	235	135
3	252	3	35	235	137
4	252	3	35	233	135
5	252	3	35	237	135

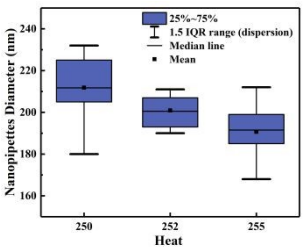

Figure 8 Box plot analysis of Heat parameter optimization.

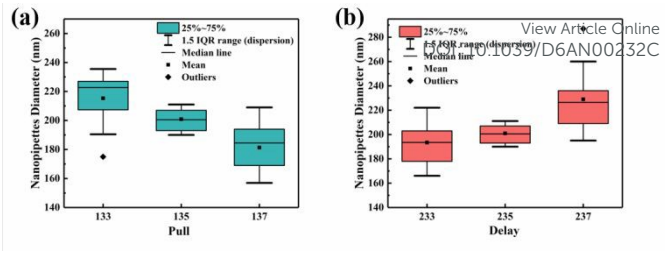

Figure 9 Box plots of nanopipette apertures tested under parameters in Table 4. (a) Groups 1, 2, 3. Effect of Pull variation on aperture (Delay fixed at 235). (b) Groups 2, 4, 5. Effect of Delay variation on aperture (Pull fixed at 135).

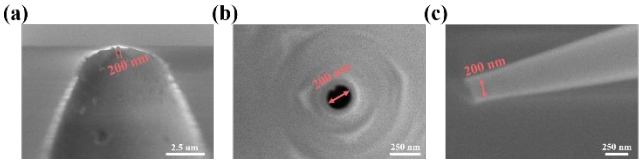

Figure 10 SEM images of a nanopipette with approximately 200 nm aperture. (a) Side view. (b) Top view. (c) Front view.

With the Heat parameter set at 252, the second critical parameter (Pull) and the third critical parameter (Delay) were optimized, with specific settings detailed in Table 4. During Pull parameter optimization (Groups 1–3), Group 2 with a Pull setting of 135 exhibited optimal stability, with a coefficient of variation (CV) of 3.83% (Figure 9a). For Delay parameter optimization, with the Pull parameter fixed at 135 (Groups 2, 4, 5), Group 4 with a Delay setting of 233 showed the lowest standard deviation (7.69 nm) and the highest Box Density Index (0.714, a value that exceeds the 0.7 threshold), thus verifying high stability of the fabrication process. Accordingly, the optimal fabrication parameters for 200 nm nanopipettes are a Heat setting of 252, a Pull setting of 135, a Delay setting of 233, a Filament setting of 3 and a Velocity setting of 35 (Figure 9b).

Nanopipettes with target aperture diameters ranging from 50 nm to 1000 nm in 100 nm increments (50, 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1000 nm) were successfully fabricated via the feature weight-guided stepwise optimization strategy. For each target size, stability validation was conducted via ten consecutive pulling operations with the corresponding optimized parameter sets. As shown in Table 5, the mean and median diameters for all eleven aperture sizes were highly consistent with their respective target values. Moreover, the SD and CV for sizes from 100 nm to 1000 nm remained at low and stable levels, whereas the 50 nm group exhibited relatively higher SD and CV values. The actual aperture diameters obtained from the ten consecutive pulls for the 50 nm target were 50 nm, 55 nm, 70 nm, 71 nm, 51 nm, 40 nm, 50 nm, 56 nm, 45 nm, and 52 nm. This variability stems from the inherent inhomogeneity of commercial borosilicate glass capillaries. Local segregation of B, Si and O is inevitable during capillary fabrication, and this segregation is aggravated in microscale drawing processes. Additionally, intrinsic microdefects (e.g., microbubbles, microcracks, and micropores) are inevitably generated during the melting and drawing of borosilicate glass. For 50 nm-scale fabrication, the geometric size of the necked region of the glass capillary is comparable to that of these microdefects. Under tensile stress, these subcritical

View Article Online
DOI: 10.1039/D6AN00232C

microdefects develop into major stress concentration sites, which directly determine the fracture position and the morphology of the nanopipette tip. After excluding the 70 nm and 71 nm outliers, the eight valid data points for the 50 nm target exhibited excellent clustering characteristics, showing a mean value of 51.0 nm (relative error < 2%), an SD of 3.5 nm, and a reduced CV of 6.9%. This result provides robust evidence that the proposed model can effectively suppress fluctuations induced by glass micro-heterogeneity, enabling high-precision control over nanopipette aperture sizes.

Excluding the two outliers of 70 nm and 71 nm, the remaining eight data points were close to the 50 nm target, demonstrating that the pulling parameters for the 50 nm nanopipette still possess good applicability and stability. The stability analysis across the eleven different nanopipette sizes confirms that the manufacturing parameters ensure good consistency and suitability for fabricating specific-sized nanopipettes. This also validates that the feature weight-guided stepwise optimization method enables the custom fabrication of nanopipettes with desired diameters. Corresponding SEM images for nanopipette orifice measurement are provided in the Supporting Information Figures S11–S21.

Notably, although the stepwise optimization strategy guided by feature weights is intuitive and highly consistent with the actual experimental workflow, from a theoretical optimization perspective it may eventually converge to a local optimum. Nevertheless, this strategy has successfully achieved stable and controllable fabrication of nanopipettes. Finally, it should be emphasized that the quantitative findings (e.g., weight values, optimal parameter sets) are specific to the Sutter P-2000 system. Nevertheless, the proposed data-driven methodology is transferable and can serve as a blueprint for optimizing other nanopipette fabrication platforms.

Table 5 Stability Analysis of Ten Consecutive Pulling Cycles for 50-1000 nm Nanopipettes. SD and CV denote standard deviation and coefficient of variation.

Target (nm)	Mean (nm)	Median (nm)	SD/nm	CV/%
50	54	51.5	10.1	18.7
100	101.2	100.5	6.4	6.28
200	199.2	200	7.3	3.65
300	294.2	288.5	19.5	6.62
400	399.8	403.5	10.6	2.66
500	507.1	505.3	15.8	3.11
600	602.9	603.9	18.7	3.1
700	692.8	692.7	22.4	3.23
800	802.6	803.7	17.6	2.19
900	898.4	898.5	23.5	2.62
1000	992.3	990.8	32.8	3.31

4 Discussion

The ANN model developed in this study, combined with three feature weight analysis methods (RF, SHAP values, Garson algorithm), achieves two key improvements. First, it overcomes the inherent limitation of RSM—incorrect characterization of nonlinear processes via linear regression or low-order polynomial fitting—by accurately fitting the nonlinear relationship between the five puller fabrication parameters and the nanopipette tip aperture. Second, it enables quantitative characterization of the feature importance of the five core puller parameters, thus breaking the dual limitations of the orthogonal experiment method in the number of analysable factors and fabrication spatial scale.

4.1 Interpretation of Key Parameter Importance

The results of parameter weight analysis show the feature importance ranking as Heat > Pull > Delay > Filament/Velocity. Among these parameters, Heat serves as the core factor for regulating nanopipette aperture, essentially because it directly modulates glass viscosity—the key variable governing glass deformation in the viscous flow state. Combined with the physical model constructed in this study, the variation of glass viscosity with temperature follows the VFT equation. By directly regulating the CO_2 laser output power, the Heat parameter controls the glass heating rate and softening degree. A higher Heat setting causes a more pronounced reduction in glass viscosity within the effective laser heating zone and greater glass softening, which in turn leads to more adequate necking deformation under subsequent tensile force and a smaller final tip aperture. On the contrary, an insufficient Heat value tends to cause inadequate glass softening, leading to fabrication failure or aperture deviation. As the fundamental parameter determining the thermal state of glass, Heat directly dominates the basic conditions for viscous flow deformation, thus becoming the primary factor affecting the aperture.

Pull and Delay are the second and third most important features in terms of feature importance. Both directly regulate the core physical conditions of the drawing process, which is based on the glass thermal state established by Heat, thus acting as key regulatory parameters for viscous flow deformation. According to the physical model, Pull adjusts the tensile stress on softened glass by controlling the solenoid current. Higher tensile stress leads to a faster deformation rate of glass in the viscous flow state, as well as a more significant reduction in aperture size. Delay controls the time interval between laser shutdown and the start of hard pulling, determining the initial temperature of glass in the cooling stage. Its value directly affects glass viscosity at the initial cooling stage, thereby altering the deformation rate and deformation sufficiency. These two parameters independently regulate the stress field and thermo-mechanical coupling time window of the drawing process. Based on the thermal input from Heat, they jointly determine the final glass deformation, thus emerging as key factors governing nanopipette aperture size.

Filament and Velocity have relatively low feature importance weights, mainly due to their narrow effective operating ranges. Velocity is only a speed threshold for triggering the hard pulling

View Article Online
DOI: 10.1039/D6AN00232C

action, acting as a "trigger condition" rather than a parameter for directly regulating deformation in the stretching process.

4.2 Synergy of Data-Driven and Physical Modeling

In this study, the data-driven model integrating ANN and feature importance analysis was combined with a physical mechanism model rooted in mechanical theory, thus forming an organic integration that enables mutual verification and synergistic optimization. This integration not only enhanced the accuracy of nanopipette aperture prediction and process parameter optimization, but also facilitated a paradigm shift in nanopipette fabrication from empirical trial-and-error to model-driven design.

The physical mechanism model effectively addresses the "black-box" limitation of the ANN model. While the ANN model can accurately capture the nonlinear input-output relationship between process parameters and aperture size, it cannot explain the intrinsic physical mechanism by which parameters regulate the aperture. In contrast, the physical model deduces the core equation for the temporal evolution of nanopipette aperture based on the basic principles of mechanics, clearly revealing the dynamic law of each parameter affecting the aperture by regulating the temperature field, viscosity field and stress field, and thus providing a clear physical mechanism interpretation for the parameter-aperture mapping relationship of the ANN model.

The data-driven model validates and optimizes the physics-based mechanistic model. The quantitative ranking of parameter feature importance from three feature weight analysis methods (RF, SHAP values, Garson algorithm) is highly consistent with the qualitative analysis of parameter regulatory effects in the physical model. This achieves complementary validation between the two models. The 135 sets of experimental parameters and corresponding aperture measurements provide sufficient data to train the ANN model within the investigated range. Nevertheless, the current dataset scale is relatively limited, and the extrapolation performance of the model beyond the studied parameter range should be treated with caution. Meanwhile**,** the high consistency between the model's predictions and experimental trends further verifies its reliability. Furthermore, the ANN model's high-precision quantitative prediction capability compensates for the quantitative deviations of the physical model induced by factors such as surface tension and laser power fluctuation. This synergy ultimately achieves the dual goals of mechanistic interpretation and accurate prediction for nanopipette aperture characterization.

The weight-guided optimization framework developed in this work allows users to adjust parameters systematically for target apertures, without relying on trial and error. It empowers researchers in fields like single-cell analysis, SECM, and drug delivery to fabricate the required nanopipettes with high precision. It also significantly reduces the time and cost of preparing experimental tools in the preliminary research stage. In subsequent work, targeted experimental data for extreme aperture sizes such as 50 nm and parameter boundary regions can be supplemented via active learning strategies, and the model can be further trained iteratively to enhance its generalization and extrapolation capabilities.

5 Conclusions

This study establishes a systematic framework for controllable, multi-scale nanopipette fabrication by integrating data-driven modelling with physical mechanism analysis. ANN accurately predicts aperture size from five strongly coupled pulling parameters (Heat, Filament, Velocity, Delay, Pull), overcoming the limitations of traditional linear or polynomial methods. RF, SHAP values and the Garson algorithm all confirm Heat as the dominant parameter, followed by Pull and Delay, providing a rational basis for hierarchical optimization. A complementary physical model derived from mechanical theory explains how these parameters govern aperture evolution and confirms the experimental trends. Guided by the parameter importance ranking, stepwise optimization successfully produces nanopipettes with 11 distinct aperture sizes from 50 nm to 1000 nm with high reproducibility. While the current framework is validated for borosilicate glass using a Sutter P-2000 puller, its transferability to other glass types or puller systems requires further investigation. This work demonstrates that combining machine learning with physical insight can transform empirical trial-and-error fabrication into a predictable, model-driven engineering process. Future research will extend this methodology to other materials (e.g., quartz) and puller systems to verify its transferability.

Author contributions

Runan Guo: methodology, investigation, data curation, visualization, writing—original draft. Zhi Chen: Software. Xue Han, Rongke Sun and Hui Lu: investigation and data curation. Yanqing Ma: funding acquisition, writing—review and editing, and supervision. Lei Ma: funding acquisition, conceptualization, supervision, methodology, validation, and writing—review and editing. All authors read and approved the final manuscript.

Conflicts of interest

There are no conflicts to declare.

Data availability

All data are available within the article and its supplementary information.

Acknowledgements

This work was financially supported by The National Key R&D Program of China (No. 2022YFC3006303), Tianjin Natural Science Foundation Project (No. 25JCZDJC01380) and the Innovation fund of Haihe Laboratory of Low-dimensional Electronic Materials (25HHDWSS00005).

Journal Name ARTICLE

View Article Online
DOI: 10.1039/D6AN00232C

References

1 S Peng, P Xia, T Wang, L Lu, P Zhang, M Zhou, F Zhao, S Hu, J Kim, J Qiu, Q Wang, X Yu and X Xu, *Acs Nano*, 2023, **17(10)**: 9543-9551.
2 J Yuan, Y Ma, M Wang, Y Zhao, M Liu, L Du and C Wu. *Advanced Materials Technologies*, 2026, **11(6)**: e01598.
3 K Kaplan, O Ulkir, and F Kuncan, *Measurement*, 2025, **256**: 117405.
4 G. Liu, L. Hao, H. Li, K. Zhang, X. Yu, D. Li, X. Zhu, D. Hao, Y. Ma and L. Ma, *Anal. Chem.*, 2022,**94**, 5248-5254.
5 Y. Takahashi, A. I. Shevchuk, P. Novak, Y. Murakami, H. Shiku, Y. Korchev and T. Matsue, *J. Am. Chem. Soc.*, 2010,**132**, 10118-26.
6 M. G. Schrlau, N. J. Dun and H. H. Bau, *ACS Nano*, 2009,**3**, 563-568.
7 Y. Ruan, H. Wang, X. Shi, Y. Xu, X. Yu, W. Zhao, H. Chen and J. Xu, *Anal. Chem.*, 2020,**93**, 1200-1208.
8 R. A. S. Nascimento, R. E. Özel, W. H. Mak, M. Mulato, B. Singaram and N. Pourmand, *Nano Lett.*, 2016,**16**, 1194-1200.
9 P. B. Tiwari, L. Astudillo, J. Miksovska, X. Wang, W. Li, Y. Darici and J. He., *Nanoscale*, 2014,**6**, 10255-10263.
10 G. Liu, W. Chen, M. Gao, l. Song, X. Hu and Y. Zhao et al., *Electrochem. Commun.*, 2018,**93**, 95-99.
11 L. Ying, A. Bruckbauer, A. M. Rothery, Y.E. Korchev and D. Klenerman, *Anal. Chem.*, 2002,**74**, 1380-1385.
12 X. Dong, Y. Huang and D. Jiang, *Anal. Chem.*, 2022,**94**, 10304-10307.
13 Y. Liu, X. Lu, Y. Peng and Q. Chen, *Anal. Chem.*, 2021,**93**, 12337-12345.
14 E. B. Tetteh, D. Valavanis, E. Daviddi, X. Xu, C. S. Santos, E. Ventosa, M. Y. Daniel, W. Schuhmann and P.R Unwin, *Angew. Chem., Int. Ed.*, 2023,**62**, e202214493.
15 F. O. Laforge, J. Carpino, S. A. Rotenberg and M. V. Mirkin, *PNAS*, 2007,**104**, 11895-11900.
16 Y. Takahashi, Y. Sasaki, T. Yoshida, K. Honda, Y. Zhou, T. Miyamoto, H. Higashi, A. Shevchuk, Y. Korchev, H. Ida, R. Hanayama and T. Fukuma, *Anal. Chem.*, 2023,**95**, 12664-12672.
17 R. Hao and B. Zhang, *Anal. Chem.*, 2015,**88**, 614-620.
18 L. Zhang, J. Liu, B. Yang, X. Kang, S. Jiang, J. Mei and C. Yang, *Microelectron. Eng.*, 2013,**103**, 131-136.
19 T. Wang, J. Liu, B. Yang, X. Chen, X. Wang and C. Yang, *J. Micromech. Microeng.*, 2016,**26**,065001.
20 O Ulkir, *Measurement*, 2025, **256**: 118089.
21 H Dong, J Lin, Y Tao, Y Jia, L Sun, W Li and H Sun. *Lab on a Chip*, 2024, **24(5)**: 1419-1440.
22 O Ulkir, F Kuncan and F D Alay. *Polymers*, **2025**, 17, 1528.
23 D. A. Dean and J. Z. Gasiorowski, *Cold Spring Harb. Protoc.*, 2011, (3)5587.
24 S. Sutter Instruments, P-2000 Micropipette Puller Operation Manual, 2020 [Z].
25 M. Wang, T. Wang, P. Cai, and X. Chen, *Small Methods*, 2019,**3**,1900025.
26 A. A. Hameed, B. Karlik and M. S. Salman, *Knowl.-Based Syst.*, 2016,**114**, 79-87.
27 S. Ranka, C. K. Mohan, K. Mehrotra and A. Menon, *Neural Netw*, 1996,**9**, 819-835.
28 M. Jahangoshai Rezaee, M. Jozmaleki and M. Valipour, *Phys. A*, 2018,**489**, 78-93.
29 L. Breiman, *Mach. Learn.*, 2001, **45**, 5-32.
30 K. J. Archer and R. V. Kimes, *Comput. Stat. Data Anal.*, 2008,**52**, 2249-2260.

View Article Online
DOI: 10.1039/D6AN00232C

Data Availability Statement

The availability of Data and Materials are manuscript and Supplementary materials.