2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

An Autonomous and Hardware-Agnostic Vision-Servoed System for Microdevice Injection

Yumin Zheng[1], Runjia Tan[2], Rui Jiao[1], Sunwoo Lee[1]

***Abstract*— Automated manipulation of nanoliter-scale implantable microdevices (IMDs) typically relies on complex, custom-built robotic setups that are difficult to reproduce and require extensive manual calibration. To address this challenge, this paper proposes an easily deployable and highly reproducible vision-servoed manipulation system for IMDs. Based on standard commercial off-the-shelf devices, the proposed platform is hardware-agnostic and eliminates the need for tedious manual calibration. The automated workflow seamlessly integrates coarse positioning, auto-focus, and marker-aided centering to achieve robust precision. The system is validated using a sub-nanoliter IMD, the microscale optoelectronic tetherless electrode (MOTE). Experimental results demonstrate that the proposed framework requires minimal manual intervention and significantly reduces operating time by 47.2 % compared to manual injection performed by an experienced user. These results pave the way for economical, high-throughput, and automated IMD-based *in vitro* and *in vivo* experiments, and beyond.**

***Index Terms*— Automation, Machine Vision, Implantable Microdevice, IMD, Manipulation, Insertion**

I. INTRODUCTION

Implantable microdevices (IMDs) have shown significant advances in recent decades in terms of size scaling (now exceeding the nanoliter limit [1]) and functionality (wireless or tetherless power and data transmission [2]–[6]). Such IMDs can enable a broad range of biomedical applications, including neural recording and stimulation, drug delivery, and chemical detection such as pH or glucose [7]–[9], all with minimal insertion damage [10]. However, a broader adoption of such IMDs has lagged a lot, not because of the qualities of IMDs but because of the lack of an automated insertion system that can provide an efficient and reliable means of inserting IMDs into targets.

To address such shortcomings, previous work has explored PEG-mediated carriers, PEG-assisted temporary bonding tools, and custom microneedle inserters [11]–[13]. Although these strategies enable placement at the target site, they introduce additional (often proprietary) materials and tools and remain largely manual; many manual techniques are also often in the realm of art rather than engineering and quite difficult to reproduce easily. Therefore, there is still a strong need for an automated and reliable alternative for a wider adoption of IMDs in practice.

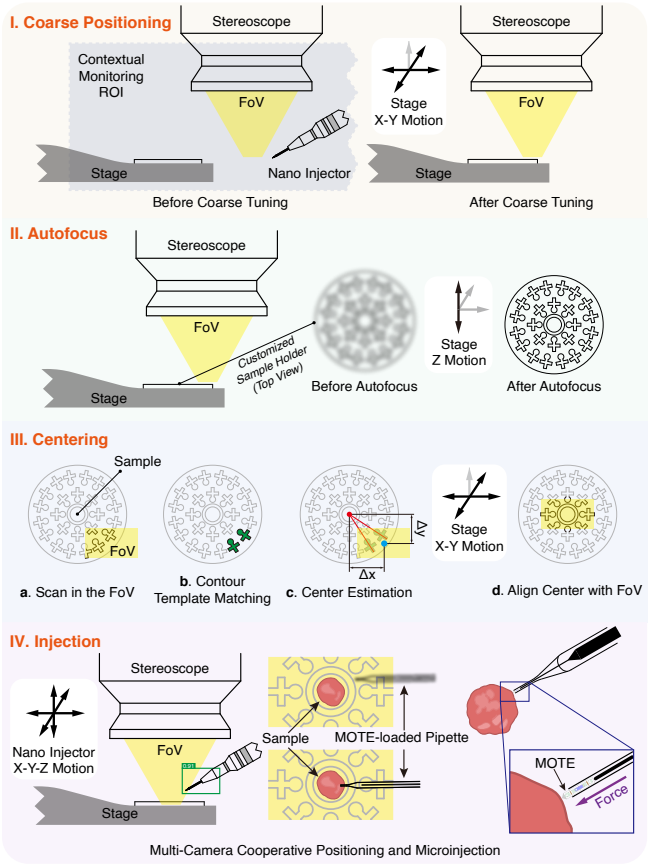


Fig. 1. Overview of the Automated Microinjection System

Although automated microinjection has been demonstrated in zebrafish embryos and larvae [14], existing systems still heavily rely on complex manual preparation and precise pre-alignment of hardware [15], [16]. Also, because payloads are typically liquid, delivery is based on direct penetration of the pipette [17], causing additional damage. In contrast, our system utilizes a self-calibrated vision-servoed approach that eliminates the need for tedious manual alignment or specialized hardware modifications. Using standard commercial off-the-shelf (COTS) precision devices, our easily deployable and highly automated system provides a "plug-and-play" solution. This is particularly valuable for studying organoids—three-dimensional multicellular assemblies

This work was supported by Nanyang Technological University, Singapore, under its Nanyang Assistant Professorship; the Ministry of Education, Singapore, under its Academic Research Funds Tier 2 (MOE-T2EP50124-0021); and the National Research Foundation, Prime Minister's Office, Singapore, under its Campus for Research Excellence and Technological Enterprise (CREATE) programme. Yumin Zheng and Runjia Tan have contributed equally. Corresponding author: Sunwoo Lee.

[1]Yumin Zheng, Rui Jiao and Sunwoo Lee are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798.(email: yumin004@e.ntu.edu.sg; jiao0038@e.ntu.edu.sg; sunwoo.lee@ntu.edu.sg)

[2]Runjia Tan is with the School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798. (email: runjia.tan@ntu.edu.sg)

where controlled, non-destructive injection is highly challenging due to their structural complexity [18].

To validate our system, we delivered a microscale optoelectronic tetherless electrode (MOTE) [5], a nanoliter-scale neural implant, on organoid dummies. Fig. 1 shows the automated injection process in four major steps. Once the operator places the sample holder on the stage (without any alignment), all subsequent motor decisions are derived from visual feedback and executed automatically. Using views from three contextual cameras whose regions of interest (ROIs) cover the entire workspace, the controller first coarsely positions the sample holder in the fixed field of view (FoV) of the camera mounted on the stereo microscope. The system then runs autofocus and uses features of the customized holder to center the sample in the image. After the sample is posed at the target location, multi-camera cooperation guides the MOTE-loaded nano injector to a non-contact pose adjacent to the sample, and the injector is then actuated to deliver the MOTE into the sample without pipette penetration. The main contributions of this work are summarized as follows:

1) A highly deployable, hardware-agnostic, one-click platform for nanoliter-scale IMD implantation built upon standard commercial precision devices.
2) Multi-camera vision pipeline—coarse positioning, autofocus, marker-aided centering, non-contact guidance—for closed-loop injection.
3) Validation in tissue-mimicking agarose with a MOTE, demonstrating a 47.2 % reduction in operating time compared to manual injection by an experienced user, while ensuring high procedural reproducibility.

The remainder of this paper is organized as follows. Section II describes the system architecture. Section III presents the automated injection approach (vision algorithms). Section IV reports on the experimental results. Section V provides a discussion and Section VI concludes the paper.

II. SYSTEM SETUP

A. Hardware Layout

All components of the system were mounted on an optical table with a 25.4 mm hole spacing, as shown in Fig. 2.

The system employs two micromanipulators (RWD MM-500) to control the three-axis motion of the injector (RWD R-480 Nano Injector) and the injection target, respectively. Due to limitations of the available API, motion commands can be sent through the RS422 interface, but the relative coordinates of each axis cannot be directly read. A custom-designed and biocompatible 3D-printed sample holder (Biomed Resin; Formlabs) is placed on the stage.

The system is highly vision-servoed, with the vision module constituting a major part of the setup. It includes three contextual monitoring cameras and a stereoscope-mounted camera. The stereoscope provides a magnification range from 7× to 110×, although it does not support motorized adjustment. To facilitate visual tracking and calibration, three ArUco markers are mounted on the optical frame, while two additional markers are mounted on the left and right micromanipulators.

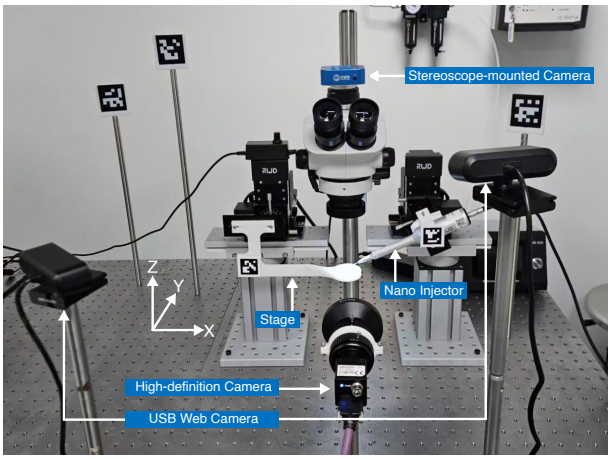
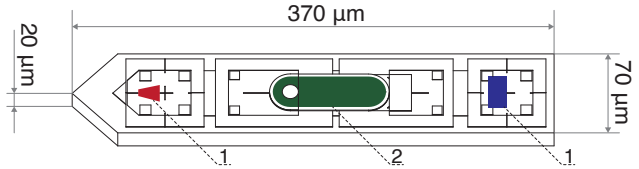


Fig. 2. System Assembly on an Optical Table.



1. Input Differential Pair 2. Photo-Voltaic Light Emitting Diode (PVLED)

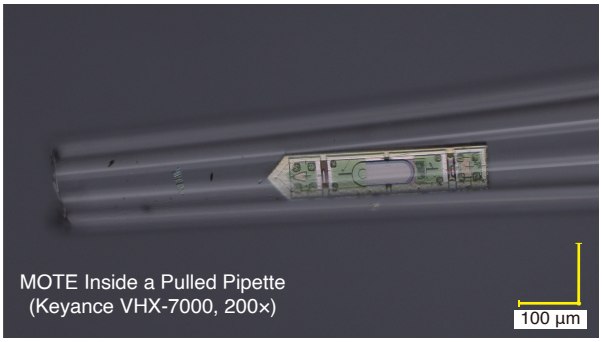


Fig. 3. Sketch of the MOTE with Key Dimensions and an Optical Microscopic Image Showing a MOTE Inside a Pulled Glass Pipette.

B. Implant Assembly

In a form factor comparable to the diameter of human hair, MOTE integrates a CMOS circuit with a photovoltaic light-emitting diode (PVLED), which absorbs incident light and transmits data via pulse-position-modulation (PPM) [2], [19]. MOTE and its derivative microdevices have demonstrated stable chronic monitoring of neural signals and pH values [4], [9]. The one-sided triangular geometry of MOTE has been optimized for facile injection.

Commercially available syringes and standard glass pipettes have inner and outer diameters that are substantially larger than those of state-of-the-art IMDs, such as a MOTE. On the other hand, pulled glass pipettes can provide a customizable and economical carrier for injection [3]. Glass pipettes with an inner diameter of approximately 100 μm are

fabricated from straight capillaries (outer diameter: 1.14 mm, inner diameter: 0.53 mm) using an RWD MP-500 pipette puller. The MOTE placement into the pulled glass pipette is conducted as follows:

1) The MOTE is immersed in isopropyl alcohol (IPA).
2) The tip of the pulled glass pipette is brought close to the MOTE, which is then drawn into the pipette by the pressure difference.
3) Residual IPA inside the pipette is removed using lint-free paper (KIMTECH Kimwipes), leaving the MOTE positioned at the pipette tip.

The MOTE placed inside the pulled glass pipette is shown in Fig. 3, along with a schematic of the MOTE and its key dimensions. The microscopy image was taken using a Keyence VHX-7000 digital microscope.

C. Sample Preparation

Since organoid culture requires a relatively long period, it is impractical to simultaneously prepare a large number of samples (dozens or even hundreds). Consequently, during repeated experiments, organoids often suffer from functional degradation, as their culture requires specific atmospheric conditions and controlled temperature and humidity, which are not incorporated in the current system. Therefore, in order to perform functional tests of the system, researchers frequently employ agarose to mimic biological tissues [20]. In our experiments, a 0.6 % agarose solution was prepared, boiled, and then cooled to allow Nanoation. After solidification, an agarose gel with a diameter of approximately 5 mm was excised and placed in the center of the sample holder.

To further approximate the visual appearance of real organoids or in vitro engineered biological tissues, we added a small amount of red dye to the sample holder to mimic the color of the cell culture medium. Typically, common culture media such as Dulbecco's Modified Eagle Medium (DMEM) or Roswell Park Memorial Institute Medium (RPMI-1640) contain phenol red, which serves as a pH indicator in the formulation [21].

III. AUTOMATIC INJECTION APPROACH

A. Coarse Positioning

In the experimental setup, the sample stage may start outside the stereoscope's FoV due to mechanical disturbances and handling constraints. The loading/unloading directly beneath the stereoscope is also impractical and risks lens damage. Consequently, the initial pose of the sample holder is placed away from the FoV, and the first automated step is the stage alignment.

Problem statement. From an arbitrary start, the stage is driven—under guidance from contextual monitoring cameras—to a predefined target point with minimal final error.

Sensing configuration. Three static ArUco markers (each 5×5 cm) are fixed to the optical table with known coordinates; two additional (dynamic) markers are attached to the moving stage. All markers are simultaneously visible to three cameras. The world origin is set at the base of the optical axis of the central camera (high-definition). The camera is approximately aligned with the stereoscope, provides sub-millimeter accuracy but exhibits low-frequency ($\sim$10 Hz) jitter under manipulator motion. To stabilize the estimate, two auxiliary cameras (USB web cameras) provide lower-resolution but higher-rate output ($\sim$30 Hz). We perform joint extrinsic/pose optimization via bundle adjustment and apply a Kalman filter for temporal denoising and sensor fusion.

The marker state is $\mathbf{x}_k = [x_k, y_k, \dot{x}_k, \dot{y}_k]^T$ (planar position and velocity). The prediction step is:

$$\mathbf{x}_{k|k-1} = \mathbf{F}\mathbf{x}_{k-1|k-1} + \mathbf{w}_{k-1}, \tag{1}$$

where $\mathbf{F}$ is the state transition matrix (e.g., constant velocity model), and $\mathbf{w}_{k-1} \sim \mathcal{N}(0, \mathbf{Q})$ is the noise of the process. The update incorporates measurements $\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k$, with $\mathbf{H}$ as a measurement matrix and $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R})$ as measurement noise. This yields real-time positions of the stage-mounted markers.

Bundle adjustment. To achieve accurate stage alignment, we employ bundle adjustment to jointly optimize the extrinsic parameters of the three cameras and the 3D poses of the ArUco markers. Bundle adjustment minimizes the reprojection error between the observed 2D marker positions in the camera images and the projected 3D marker positions based on the estimated camera parameters and marker poses. For N markers observed by M cameras, let $\mathbf{p}_{ij} = [u_{ij}, v_{ij}]^T$ denote the observed 2D image coordinates of marker i in camera j, and $\hat{\mathbf{p}}_{ij}(\mathbf{X}_i, \mathbf{C}_j)$ denote the projected 2D coordinates of the 3D marker position $\mathbf{X}_i = [X_i, Y_i, Z_i]^T$ using the camera parameters $\mathbf{C}_j$ (including intrinsic and extrinsic parameters). The reprojection error is defined as:

$$\mathbf{e}_{ij} = \mathbf{p}_{ij} - \hat{\mathbf{p}}_{ij}(\mathbf{X}_i, \mathbf{C}_j). \tag{2}$$

The bundle adjustment problem is formulated as a nonlinear least-squares optimization:

$$\min_{\{\mathbf{X}_i\},\{\mathbf{C}_j\}} \sum_{i=1}^{N}\sum_{j=1}^{M} w_{ij}\|\mathbf{e}_{ij}\|^2, \tag{3}$$

where w_{ij} is a binary weight that indicates whether the marker i is visible on the camera j. The optimization is solved iteratively using a method such as Levenberg-Marquardt, refining the estimates of $\mathbf{X}_i$ and $\mathbf{C}_j$. This process ensures that the static markers' known coordinates are consistent with the dynamic markers' estimated poses, improving the accuracy of the stage's position estimate.

The filtered positions from the Kalman filter, combined with the optimized marker poses from bundle adjustment, drive a trajectory planner that brings the stage to the target. Once within $\sim$0.5 cm (macro-alignment), a fine alignment routine centers the ROI within the stereoscope's FoV.

B. Autofocus

Autofocus is a critical pre-centering step. For a grayscale image $I(x, y)$ captured in axial position z, the Laplacian response is $L(x, y; z) = (K * I)(x, y; z)$, where $*$ denotes

the 2-D convolution (symmetric padding). We use the 8-neighborhood kernel:

$$K = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}. \quad (4)$$

The focus score is the sample variance of the Laplacian values $\{L_i\}_{i=1}^{N}$: $S(z) = \frac{1}{N-1}\sum_{i=1}^{N}(L_i - \bar{L})^2$, $\bar{L} = \frac{1}{N}\sum_{i=1}^{N} L_i$, with N pixels. Since defocus attenuates high spatial frequencies, $S(z)$ is typically maximized near the best focus and is locally unimodal.

We maximize $S(z)$ along the optical axis using a ternary search on $[l, h] \subseteq [z_{\min}, z_{\max}]$: at each iteration set $m_1 = l+(h-l)/3$ and $m_2 = h-(h-l)/3$, evaluate $S(m_1), S(m_2)$, and update $(l, h) \leftarrow (m_1, h)$ if $S(m_1) < S(m_2)$, otherwise $(l, h) \leftarrow (l, m_2)$. Stop when $|h - l| \le \varepsilon$ and return $\hat{z} = \arg\max_{z\in\{l,m_1,m_2,h\}} S(z)$.

C. Marker-aided Centering

For steroscopes without motorized magnification control, achieving both wide field-of-view (FoV) and sharp image quality is a challenge. We address the issue through a specially designed sample holder that enables precise localization, even within a narrow FoV, thereby ensuring clear image capture. The sample holder consists of a 3.2 cm-diameter disk with a 7 mm cylindrical cavity at the center to accommodate the sample. Twenty-four markers are arranged at different angles around the holder to assist in localization, as shown in Fig. 4. The sample holder is 3D-printed using a biocompatible resin (Formlabs Form 4B, BioMed White), and the marker regions are thinned and perforated to improve the stability of pattern recognition.

The markers combine cross and circular primitives, which are robust for localization: crosses provide directional cues, whereas circles offer isotropy. However, circles lack directionality, and crosses are directionally ambiguous (two-fold). We therefore design a hybrid marker in which each cross (4 mm arm length, 1 mm width) is paired with a 1 mm-radius circle placed at one end of the cross and oriented toward the sample-holder center. This asymmetry shifts the pattern's centroid toward the circular end and produces a unique inward-pointing vector, allowing unambiguous determination of where the center of the sample-holder lies relative to the detected marker.

To ensure robust center estimation, additional complete markers are required for recognition. Once any edge of the sample holder enters the stereoscope's FoV, a 5 × 5 grid search is performed on the X–Y plane, moving the stage in a range of 1.1 cm relative to the current position. The stitching process is implemented using the Very Fast Sequential Micrograph Stitching (VFSMS) algorithm [22], which incrementally aligns and merges images based on sequential pairwise registration. The process can be mathematically described as follows:

Sequential Registration. For each pair of consecutive images I_t and I_{t+1}, the transformation $T_{t,t+1}$ is estimated using

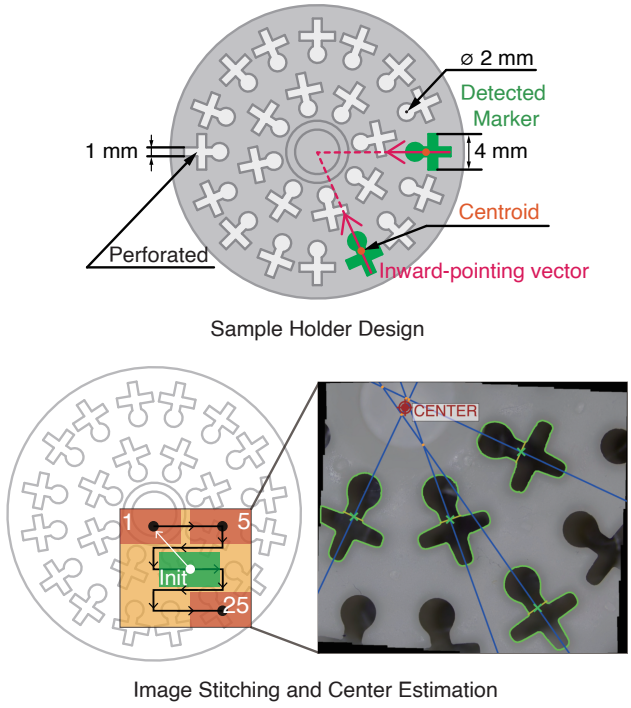


Fig. 4. Customized Markers For Image Stitching and Center Estimation.

feature matching and incremental homography:

$$T_{t,t+1} = \arg\min_{T} \sum_{(x,y)\in\mathcal{M}} \|I_t(x,y) - I_{t+1}(T(x,y))\|^2, \quad (5)$$

where $\mathcal{M}$ is the set of matching points, and T is a 2D homography matrix.

Cumulative Transformation. The global transformation $T_{1,t}$ for image I_t relative to the first image I_1 is computed incrementally:

$$T_{1,t} = T_{1,t-1} \circ T_{t-1,t}, \quad (6)$$

where $\circ$ denotes composition of homographies.

Blending. The final stitched image I_{stitched} is a weighted blend of all transformed images to reduce seams:

$$I_{\text{stitched}}(x,y) = \sum_{t=1}^{N} w_t(x,y) \cdot I_t(T_{1,t}^{-1}(x,y)), \quad (7)$$

where $w_t(x, y)$ is a weight function (e.g., based on overlap and distance from seam). The process of scanning stitching is shown in Fig. 4. After image stitching, we perform contour-based template matching against a rotated cross-shaped template. Let $\mathcal{C}$ be a detected target contour and $\mathcal{T}$ be the template cross contour. For a rotation angle $\theta \in \Theta$ (e.g., we set $0^\circ : 1^\circ : 179^\circ$), define the rotated template $\mathcal{T}_\theta = \mathbf{R}_\theta\mathcal{T}$, where $\mathbf{R}_\theta$ is the 2-D rotation. For a binary region $\Omega \subset \mathbb{R}^2$ with indicator $\mathbf{1}_\Omega(x, y)$, its raw and central moments are:

$$m_{pq} = \sum_{(x,y)\in\Omega} x^p y^q, \mu_{pq} = \sum_{(x,y)\in\Omega} (x-\bar{x})^p (y-\bar{y})^q, \quad (8)$$

with centroid $\bar{x} = \frac{m_{10}}{m_{00}}$, $\bar{y} = \frac{m_{01}}{m_{00}}$ and normalized moments $\eta_{pq} = \mu_{pq}/\mu_{00}^{1+\frac{p+q}{2}}$. The seven Hu invariant moments $\{\phi_k\}_{k=1}^{7}$ are fixed algebraic combinations of η_{pq} and are rotation/scale/translation invariant [23], [24]. Define the Hu-moment dissimilarity between contours as:

$$D_{\mathrm{Hu}}(\mathcal{C}, \mathcal{T}_\theta) \;=\; \sum_{k=1}^{7} \left| \log\left|\phi_k(\mathcal{C})\right| - \log\left|\phi_k(\mathcal{T}_\theta)\right| \right|. \qquad (9)$$

We accept a match $(\mathcal{C}, \theta)$ if $D_{\mathrm{Hu}}(\mathcal{C}, \mathcal{T}_\theta) < \tau$ (e.g., we set $\tau = 0.1$). For each accepted contour $\mathcal{C}$, we apply an Intersection-over-Union-based (IoU) non-maximum suppression (NMS), yielding a set of non-overlapping instances. An orthogonal-line extraction for non-overlapping contours to determine the inward-pointing vector, the pseudo code is shown in Algorithm 1.

Algorithm 1 Orthogonal-Line Extraction per Contour

for all $\mathcal{C}$ **do**
  $c \leftarrow (m_{10}/m_{00},\, m_{01}/m_{00})$
  $L_{\min} \leftarrow +\infty, \quad \alpha^\star \leftarrow$ null
  for all $\alpha \in \mathcal{A}$ **do**
    $u \leftarrow (\cos\alpha, \sin\alpha)$
    $\ell_\alpha \leftarrow \mathrm{Line}(c, u)$
    $(p^-, p^+) \leftarrow \textsc{ClosestPair}(\ell_\alpha \cap P(\mathcal{C}),\, c)$
    if p^- and p^+ exist **then**
      $L \leftarrow \|p^+ - p^-\|_2$
      if $L < L_{\min}$ **then**
        $L_{\min} \leftarrow L, \quad \alpha^\star \leftarrow \alpha$
      end if
    end if
  end for
  $u_\perp \leftarrow (-\sin\alpha^\star, \cos\alpha^\star)$
  $(a, b) \leftarrow u_\perp; \quad c_0 \leftarrow -(a\,c_x + b\,c_y)$
  append line (a, b, c_0) to $\mathcal{L}$
end for
return $\mathcal{L}$

Given lines $\tilde{\ell}_i : \; a_i x + b_i y + c_i = 0$ with unit normals (a_i, b_i), we estimate the center as the (weighted) least-squares minimizer $\hat{c} = (\hat{x}, \hat{y}) = \arg\min_{x,y} \sum_i w_i\,(a_i x + b_i y + c_i)^2$, with $w_i \equiv 1$. $\hat{c}$, defines the translation of the sample holder and is solved by:

$$\begin{bmatrix} \sum_i a_i^2 & \sum_i a_i b_i \\ \sum_i a_i b_i & \sum_i b_i^2 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = - \begin{bmatrix} \sum_i a_i c_i \\ \sum_i b_i c_i \end{bmatrix}. \qquad (10)$$

D. Injection

The final injection stage comprises coarse and fine alignment. Because mounting the glass pipette introduces placement error, coarse alignment cannot rely on fixed target coordinates; instead, we localize the pipette tip in the image. We trained a Grounding DINO detector [25] in three categories: (i) nano injector with a mounted pipette during coarse alignment, (ii) agarose gel samples during fine alignment, and (iii) pipettes during fine alignment—each with ∼100 images. The detector's bounding boxes then seed the Segment Anything Model (SAM) [26] to obtain instance masks.

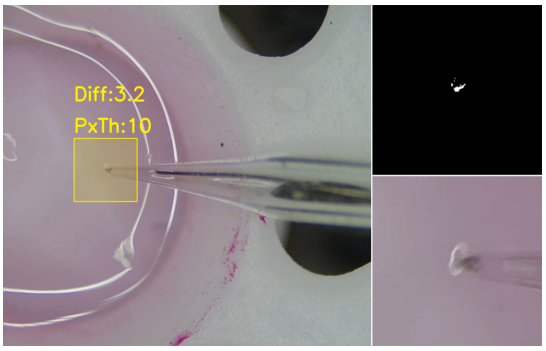


Fig. 5. Schematic Diagram of the Frame Difference Method. Left: Full-scale view with the marked ROI. Right: Frame difference and rescaled ROI image.

The injection procedure begins with online tip segmentation during *coarse* alignment. We obtain a pipette mask by running Grounding DINO detections followed by SAM segmentation, maintain the estimate for ∼1 s to stabilize and reject outliers, and then estimate the tip as the mean of ten samples near the mask's bottom-left extremity. We compute the pixel offset between the estimated tip and the desired tip location and, using the camera's intrinsic parameters, convert this 2D offset into the required metric displacement.

During fine alignment, the pipette initially appears defocused. We first localize the agarose gel sample with Grounding DINO and translate the sample along the X-axis until it enters the field of view at the margin to reduce visual interference with subsequent tip detection. The glass pipette is then advanced along the Z-axis toward the stage to improve focus until the tip is reliably detected. Using the same estimator as in coarse alignment, we obtain the tip position and compute the required X-axis correction to align the tip with the gel region. A 200×200 px region of interest (ROI) is placed at the target location for motion detection via the Frame Difference Method (FDM) [27] (Fig. 5). FDM computes absolute inter-frame intensity differences; slight contact with the gel produces pronounced surface deformation and a large difference signal, which we use to stop motion and place the tip precisely at the contact point. The nano injector is then activated to deliver the MOTE chip.

IV. RESULT

A. Trajectory

To validate the accuracy of the coarse-positioning module, we commanded a predefined trajectory and instructed the left micromanipulator (equipped with the sample holder) to follow it while recording from three synchronized cameras. The manipulator's step size was 0.924 µm; along each Cartesian direction, we executed 5000 steps (total commanded travel per axis: 5000 × 0.924 µm = 4.62 mm) and then returned to the origin.

We estimated the 3-D pose using two pipelines: (i) a baseline Perspective-n-Point (PnP) solution from ArUco marker detections, and (ii) a Kalman-filtered, bundle-adjusted (KF + BA) fusion that aggregates detections over time, with bundle

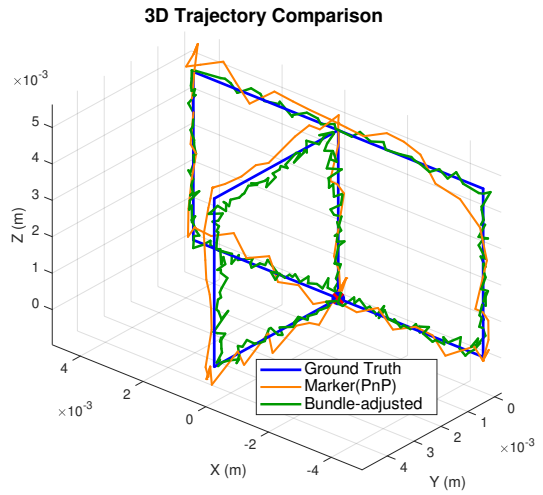


Fig. 6. Trajectory of the Ground Truth, PnP estimate, and Bundle-adjusted estimate.

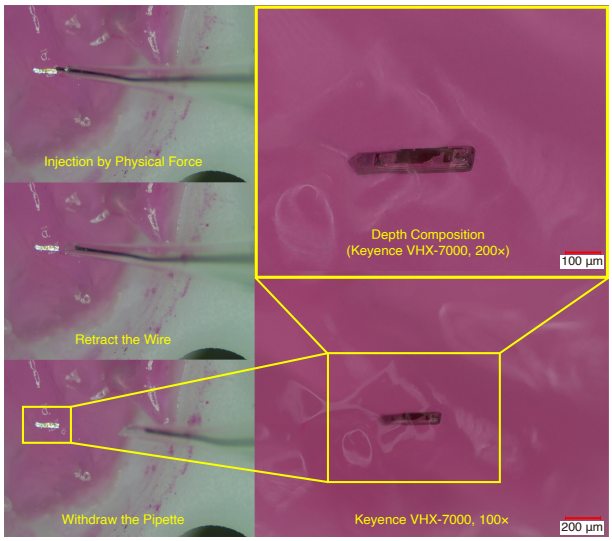


Fig. 7. Example Automated Injection Process of a MOTE.

adjustment detailed in Sec. III-A. To obtain per-axis error statistics, the trajectory was designed so that only one axis moved at a time, isolating accuracy on the actuated axis and cross-axis jitter on the others. We then visualized the ideal trajectory and the two estimates (as shown in Fig. 6. Compared to PnP, KF + BA reduces the per-axis error by 21.2 % and the jitter from 10.5 % to 2.7 %, thus mitigating control-signal chatter during stage movement.

B. Experiment on MOTE

We performed 30 automated injection trials using two MOTEs that were alternated and reused after each injection. As shown in Table I, it reduces the operating time by 47.2 % compared to the manual operation by an experienced user [3]. All automated trials were performed on a workstation equipped with an AMD Ryzen 7 7700X CPU and an NVIDIA GeForce RTX 4080 Super GPU.

TABLE I
AVERAGE TIME FOR EACH STAGE IN THE AUTOMATIC INJECTION

Phase	Stage Alignment	Autofocus	Centering	Injection
Average Time (s)	30.5	24.2	56.3	47.1

Fig. 7 illustrates an example of an automated MOTE-injection process. The sequence shows the MOTE being advanced from a pulled glass micropipette by a silver wire and inserted into agarose gel, followed by the retraction of the silver wire under the control of a nanoinjector and finally the withdrawal of the micropipette by a micro-manipulator. On the right, two microscopic views are provided at 100× and 200× magnification, respectively. The 200× image was acquired using the depth-up function of a Keyence digital microscope, which composites multiple focal planes to eliminate defocusing artifacts caused by targets located at different depths.

After injection, the agarose gel containing the implanted MOTE was examined under a microscope. The MOTE was optically powered by a Thorlabs SOLIS-623C light source, and its emitted pulses were detected with a Thorlabs APD410A photodetector and digitized using a PicoScope. Since the agarose gel itself does not carry neural or information-bearing signals, the recorded PPM waveform does not convey meaningful data. However, the presence of expected PPM pulses confirms that the implanted MOTE remains functional, as illustrated in Fig. 8.

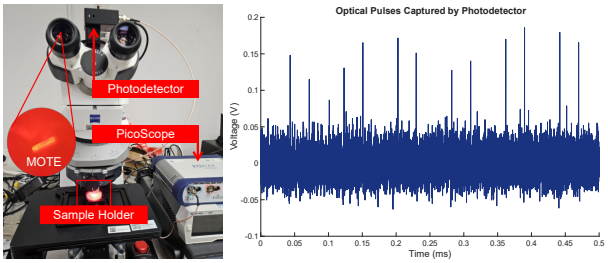


Fig. 8. Functional Verification of the MOTE in Agarose Gel.

V. DISCUSSION

Our system should be able to accommodate any implantable microdevices (IMDs) with a cross-sectional diameter on the order of, or smaller, than 0.53 mm. In addition to a MOTE, we were also able to inject a 008004 surface mount capacitor (250 μm × 125 μm × 125 μm, muRata), with a pulled pipette with a matching inner diameter (Fig. 9). In both cases, we have demonstrated that our automated injection approach outperforms even a skilled user injecting manually. While the IMD placement into a pulled micropipette is not yet fully automatic, they can also be automated in a similar manner; it is likely that, in practice, pipettes pre-loaded with the IMDs will be batch-prepared and supplied to users.

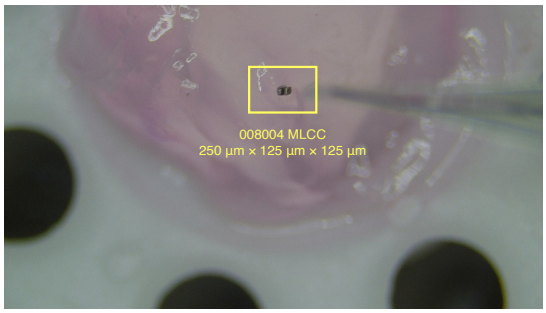


Fig. 9. Validation of System Scalability for Microdevices of Comparable Dimensions to a MOTE.

VI. CONCLUSIONS

Although recent advances in IMDs have achieved remarkable scaling and added functionality, their wider adoption in complex biological studies, such as those involving organoids, has been hindered by the lack of a standardized and accessible insertion methodology. Our easy-to-build and highly reproducible vision-servoed system bridges this gap by providing a calibration-free workflow that transforms a traditionally expert-dependent task into a robust, automated engineering process, and is particularly well suited for the automated injection of nanoliter-scale IMDs.

Our system achieves higher reliability and efficiency than the manual procedures performed by experienced operators, reducing the operation time by 47.2 %. Moreover, our architecture is hardware-agnostic; by utilizing vision algorithms to compensate for mechanical variance, the system maintains high precision across various brands of commercial micromanipulators and imaging devices. By eliminating the need for tedious manual calibration and specialized training, this platform offers a "plug-and-play" solution that is easily portable to any standard biomedical laboratory. This work provides a critical engineering solution that facilitates the broader adoption of IMDs, enabling high-throughput biological studies previously difficult to conduct.

REFERENCES

[1] A. C. Molnar, S. Lee, A. Cortese, P. McEuen, S. Sadeghi, and S. Ghajari, "Nanoliter-scale autonomous electronics: Advances, challenges, and opportunities," in *2021 IEEE Custom Integrated Circuits Conference (CICC)*, 2021, pp. 1–6.

[2] S. Lee, A. J. Cortese, A. P. Gandhi, E. R. Agger, P. L. McEuen, and A. C. Molnar, "A 250 μm × 57 μm microscale opto-electronically transduced electrodes (motes) for neural recording," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 6, pp. 1256–1266, 2018.

[3] S. Lee, A. J. Cortese, A. Mok, C. Wu, T. Wang, J. U. Park, C. Smart, S. Ghajari, D. Khilwani, S. Sadeghi, Y. Ji, J. H. Goldberg, C. Xu, P. L. McEuen, and A. C. Molnar, "Fabrication of injectable micro-scale opto- electronically transduced electrodes (motes) for physiological monitoring," *Journal of Microelectromechanical Systems*, vol. 29, no. 5, pp. 720–726, 2020.

[4] R. Jiao, Y. Zheng, S. Ghajari, S. Sadeghi, A. J. Cortese, P. L. McEuen, A. C. Molnar, and S. Lee, "A 43 μm × 269 μm, light-tolerant and power-adaptive forward-bulk optoelectrical microsystem for tetherless neural recording," in *2025 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2025, pp. 1–5.

[5] S. Lee, S. Ghajari, S. Sadeghi, Y. Zheng, H. Zahr, A. J. Cortese, W. Gu, K. Choe, A. Mok, M. Wallace, R. Jiao, C. Wu, J. C. Werth, W. Fan, P. Mogalipuvvu, J. Lammerding, T. Wang, J. H. Goldberg, P. L. McEuen, C. Xu, and A. C. Molnar, "A subnanolitre tetherless optoelectronic microsystem for chronic neural recording in awake mice," *Nature Electronics*, vol. 8, no. 12, pp. 1259–1271, 2025. [Online]. Available: https://doi.org/10.1038/s41928-025-01484-1

[6] R. Jiao, Y. Zheng, A. M. Nibras, S. Ghajari, S. Sadeghi, A. J. Cortese, P. L. McEuen, A. C. Molnar, and S. Lee, "A 43 μm × 269 μm light-adaptive optoelectronic autonomous microsystem for neural recording," *IEEE Transactions on Biomedical Circuits and Systems*, pp. 1–14, 2026.

[7] H. Kim, B. Rigo, G. Wong, Y. J. Lee, and W.-H. Yeo, "Advances in wireless, batteryless, implantable electronics for real-time, continuous physiological monitoring," *Nano-Micro Letters*, vol. 16, no. 1, 2023.

[8] S. Oh, J. Jekal, J. Liu, J. Kim, J.-U. Park, T. Lee, and K.-I. Jang, "Bioelectronic implantable devices for physiological signal recording and closed-loop neuromodulation," *Advanced Functional Materials*, vol. 34, no. 41, p. 2403562, 2024.

[9] S. Ghajari, S. Lee, S. L. Norris, P. L. McEuen, and A. C. Molnar, "Redox-enabled microscale opto-electronically transduced electrodes (remotes)," in *2023 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2023, pp. 1–5.

[10] A. Khalifa, S. Lee, A. C. Molnar, and S. Cash, "Injectable wireless microdevices: challenges and opportunities," *Bioelectronic Medicine*, vol. 7, no. 1, p. 19, 2021. [Online]. Available: https://doi.org/10.1186/s42234-021-00080-w

[11] S. A. Sigurdsson, Z. Yu, J. Lee, and A. Nurmikko, "A method for large-scale implantation of 3d microdevice ensembles into brain and soft tissue," *Microsystems & Nanoengineering*, vol. 6, p. 97, 2020. [Online]. Available: https://doi.org/10.1038/s41378-020-00210-5

[12] A. Khalifa, A. Eisape, B. Coughlin, and S. Cash, "A simple method for implanting free-floating microdevices into the nervous tissue," *Journal of Neural Engineering*, vol. 18, no. 4, p. 045004, 2021.

[13] A. J. Cortese, C. L. Smart, T. Wang, M. F. Reynolds, S. L. Norris, Y. Ji, S. Lee, A. Mok, C. Wu, F. Xia, N. I. Ellis, A. C. Molnar, C. Xu, and P. L. McEuen, "Microscopic sensors using optical wireless integrated circuits," *Proceedings of the National Academy of Sciences*, vol. 117, no. 17, pp. 9173–9179, 2020. [Online]. Available: https://www-pnas-org.remotexs.ntu.edu.sg/doi/abs/10.1073/pnas.1919677117

[14] Y. Zhao, H. Sun, X. Sha, L. Gu, Z. Zhan, and W. J. Li, "A review of automated microinjection of zebrafish embryos," *Micromachines*, vol. 10, no. 1, 2019. [Online]. Available: https://www.mdpi.com/2072-666X/10/1/7

[15] Z. Guo, N. Ai, W. Ge, and Q. Xu, "Design of an automated robotic microinjection system for batch injection of zebrafish embryos and larvae," *Microsystems Nanoengineering*, vol. 10, no. 1, p. 20, 2024. [Online]. Available: https://doi.org/10.1038/s41378-023-00645-6

[16] C. Qian, X. Yu, M. Tong, S. Zhuang, and W. Lin, "Visual-guided solutions in automated zebrafish larva heart micro-injection," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1395–1402, 2022.

[17] W. Wang, X. Liu, D. Gelinas, B. Ciruna, and Y. Sun, "A fully automated robotic system for microinjection of zebrafish embryos," *PLOS ONE*, vol. 2, no. 9, pp. 1–7, 09 2007. [Online]. Available: https://doi.org/10.1371/journal.pone.0000862

[18] O. L. Eichmüller and J. A. Knoblich, "Human cerebral organoids — a new tool for clinical neurology research," *Nature Reviews Neurology*, vol. 18, no. 11, pp. 661–680, 2022. [Online]. Available: https://doi.org/10.1038/s41582-022-00723-9

[19] S. Lee, A. J. Cortese, P. Trexel, E. R. Agger, P. L. McEuen, and A. C. Molnar, "A 330m×90m opto-electronically integrated wireless system-on-chip for recording of neural activities," in *2018 IEEE International Solid-State Circuits Conference - (ISSCC)*, 2018, pp. 292–294.

[20] W. M. A. M. T. K. S. S. S. S. M. A. Y. Mustari Afrina, Nishidate Izumi, "Agarose-based tissue mimicking optical phantoms for diffuse reflectance spectroscopy," *JoVE*, no. 138, p. e57578, 2018.

[21] S. Weiskirchen, S. K. Schröder, E. M. Buhl, and R. Weiskirchen, "A beginner's guide to cell culture: Practical advice for preventing needless problems," *Cells*, vol. 12, no. 5, 2023. [Online]. Available: https://www.mdpi.com/2073-4409/12/5/682

[22] B. Ma, X. Ban, H. Huang, W. Liu, C. Liu, D. Wu, and Y. Zhi, "A fast algorithm for material image sequential stitching," *Computational Materials Science*, vol. 158, pp. 1–13, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0927025618307158

[23] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, 1962.
[24] Y. Zhang, Y. Hao, Q. Li, and H. Yin, "Improved hu moment invariant empowered fast image contour registration algorithm," in *2025 5th International Conference on Neural Networks, Information and Communication Engineering (NNICE)*, 2025, pp. 403–406.
[25] S. Liu, Z. Zeng, T. Ren, F. Li, H. Zhang, J. Yang, C. Li, J. Yang, H. Su, J. Zhu *et al.*, "Grounding dino: Marrying dino with grounded pre-training for open-set object detection," *arXiv preprint arXiv:2303.05499*, 2023.
[26] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo *et al.*, "Segment anything," *arXiv preprint arXiv:2304.02643*, 2023.
[27] R. Jain and H.-H. Nagel, "Difference and accumulative difference pictures in dynamic scene analysis," in *Image and Vision Computing*, vol. 2, no. 2. Elsevier, 1981, pp. 99–108.