RESEARCH ARTICLE | JULY 16 2026

Optoelectronic sensors with photonic memory for trajectory prediction

Fedor L. Ivanov ; Viktor V. Krasnikov ; Andrey A. Grunin ; Artem V. Chetvertukhin ;
Andrey A. Fedyanin

APL Mach. Learn. 4, 036102 (2026)
https://doi.org/10.1063/5.0335025

APL Machine Learning ARTICLE pubs.aip.org/aip/aml

Optoelectronic sensors with photonic memory for trajectory prediction

Cite as: APL Mach. Learn. 4, 036102 (2026); doi: 10.1063/5.0335025
Submitted: 21 March 2026 • Accepted: 30 June 2026 •
Published Online: 16 July 2026

View Online Export Citation CrossMark

Fedor L. Ivanov, Viktor V. Krasnikov, Andrey A. Grunin, Artem V. Chetvertukhin, and Andrey A. Fedyanin[a)]

AFFILIATIONS

Faculty of Physics, Lomonosov Moscow State University, Moscow 119991, Russia

[a)]Author to whom correspondence should be addressed: fedyanin@nanolab.phys.msu.ru

ABSTRACT

Neuromorphic vision systems, inspired by biological vision, offer high energy efficiency and sensor-level data processing. Among these systems, neuromorphic optoelectronic sensors are particularly promising because they exhibit biomimetic responses to incident light. A key feature of such devices is an inherent memory-like response, hereafter referred to as "photonic memory": sensor conductance depends not only on instantaneous illumination but also on exposure history, enabling temporal integration of information. Here, we study how photonic memory impacts trajectory prediction by comparing a ZnO-based photonic-memory sensor with a conventional image sensor in an experimentally grounded numerical framework. We experimentally characterize the non-volatile conductivity dynamics of the synapse under controlled illumination and darkness and develop a compact mathematical model of its memory behavior. The resulting imaging model is further validated using experimentally reconstructed motion patterns based on measured ZnO-based synaptic-pixel responses. We then evaluate both sensors in a trajectory prediction task across different noise levels and numbers of input frames supplied to an artificial neural network. Our results show that, within this experimentally grounded numerical benchmark, photonic memory provides a clear advantage in the single-frame regime, where the conventional sensor can rely only on spatial priors rather than explicit temporal information. However, as the number of input frames increases, this advantage diminishes and is eventually lost. A conventional image sensor then achieves superior multi-frame performance. These findings suggest that photonic memory alone is insufficient for optimal multi-frame trajectory prediction and must be complemented by additional biologically inspired mechanisms, such as synaptic-like depression implemented at the level of device physics and sensor architecture. More broadly, our work outlines design directions for optoelectronic vision hardware that exploit a richer and more versatile space of device-level temporal responses.

I. INTRODUCTION

Artificial vision technologies have achieved significant breakthroughs and become essential components of a wide range of intelligent systems, from autonomous vehicles[1] to advanced robotics.[2] Unlike biological vision, which incorporates substantial pre-processing of visual information,[3] most conventional artificial vision systems are based on the von Neumann architecture.[4,5] A schematic diagram in Fig. 1(a) illustrates the architecture of a conventional vision system, where optical signals are first captured by a sensor and subsequently converted into electrical signals. These signals are then transmitted to a separate processing unit, where both pre- and post-processing are performed. However, due to the physical decoupling of sensing, memory, and computation, the system requires continuous data shuttling between the processing unit and the memory module. This inefficient data transfer introduces a fundamental limitation known as the von Neumann bottleneck,[6,7] which restricts processing speed and leads to latency and bandwidth limitations, particularly in high-speed machine vision applications.[8]

Neuromorphic computing, unlike conventional von Neumann architectures, is designed to emulate biological information processing, enabling more energy-efficient computation.[9] This paradigm is particularly promising for applications in healthcare and robotics,[10–12] where low-latency and low-power processing are critical. By co-locating sensing, memory, and computation, neuromorphic systems help overcome the architectural limitations of von Neumann-based hardware,[4] while better matching

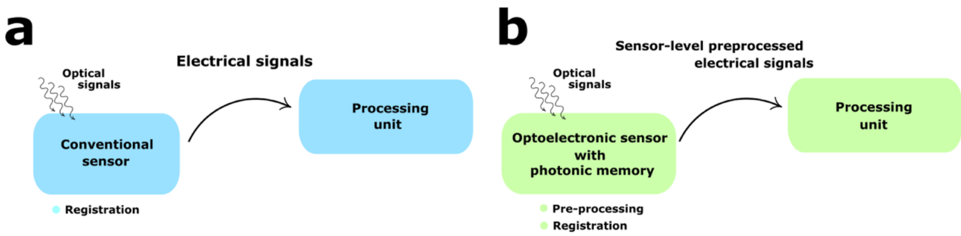

FIG. 1. Schematic representations of artificial vision systems: (a) a conventional image-sensing architecture, in which optical input is registered by the sensor and transferred as raw electrical signals to a separate processing unit; (b) a photonic-memory sensing architecture, in which registration and sensor-level pre-processing occur within the memory-enabled sensor before the signal is transferred to the processing unit.

the scalability and adaptivity requirements of future intelligent systems.[13]

Biological neural computation relies on the interaction of neurons, which communicate primarily through synapses acting as intercellular junctions for signal transmission.[14] Recent advances in neuromorphic devices have led to the development of optoelectronic structures[15] that mimic synaptic properties by responding to electromagnetic radiation and altering their electrical characteristics accordingly.[16–24] Building on these structures, researchers have demonstrated optoelectronic synaptic devices that not only directly respond to optical stimuli but also provide temporary data storage,[23] in-sensor pre-processing,[25] and real-time processing of visual information.[8] Recent studies have further demonstrated that illumination-history-dependent material dynamics can be used not only for sensing, but also for direct information processing. For example, light-responsive organic matter has been used for in-material reservoir computing, where optically induced molecular dynamics provide adaptive material-level memory and processing functionality.[26] Similarly, InSe-based dynamic memristors have been demonstrated as optoelectronic reservoir-computing elements for temporal and spatial signal-recognition tasks.[27] These works show that light-driven material and device dynamics can serve as computational resources for neuromorphic processing. Unlike these studies, the present work focuses on a complementary functionality: we use experimentally measured ZnO photonic-memory dynamics as a pixel-level sensor response and evaluate how such sensor-level temporal pre-processing affects trajectory prediction when coupled to a CNN–LSTM pipeline. Figure 1(b) illustrates a schematic artificial vision system that employs this neuromorphic approach, in which optical signals are registered and pre-processed directly within the photonic-memory sensor before being transferred to the processing unit, thereby reducing the data-transfer bottleneck and improving overall system efficiency.

The memory properties of optoelectronic synaptic devices, including long-term retention and plasticity, have attracted significant scientific interest.[28–32] Among these, non-volatile photonic memory is particularly notable, as it allows optically induced states to persist even after the removal of external stimuli.[28,33] In such devices, the response to optical excitation can be sustained for timescales significantly longer than the duration of the incident radiation, which makes non-volatile photonic memory especially promising for neuromorphic computing and optical information storage.

In particular, such residual photoconductivity plays a key role in information retention. This phenomenon is especially beneficial for applications requiring efficient recognition of motion direction, as demonstrated in Ref. 34. The ability to maintain optically induced electrical states without continuous energy input makes non-volatile photonic memory a promising approach for sensory pre-processing. In the following discussion, we refer to this delayed and persistent response to optical exposure as photonic memory.

In this paper, we develop an experimentally grounded mathematical model of a ZnO-based photonic-memory sensor and present a comparative analysis with a conventional image sensor within a numerical trajectory-prediction framework. Conventional image sensors based on semiconductor technology are mainly classified into charge-coupled devices (CCD) and complementary metal–oxide–semiconductor (CMOS) active pixel sensors.[8] These sensors operate as two-dimensional (2D) arrays that capture information from the external environment as digital images, relying on a memory unit for data storage and a processing unit that executes computer vision algorithms for tasks such as object recognition and tracking. Their operation is governed by externally controlled timing and digital signals, acquiring visual data as a sequence of discrete "snapshot" frames. This temporal quantization of visual information, dictated by a predefined frame rate, does not correspond to the continuous dynamics of real-world motion. Each recorded frame conveys pixel information uniformly, regardless of whether any changes have occurred since the previous frame was captured.[35]

In contrast, the photonic-memory sensor considered in this work differs from a conventional image sensor only in its 2D pixel array, which is composed of optoelectronic synaptic devices. This modification is incorporated into the mathematical model to capture photonic-memory effects, whereby optically induced states are retained beyond the period of external optical stimulation. Unlike conventional sensors, which rely solely on instantaneous signal acquisition, the presence of photonic memory enables temporally extended retention and in-sensor processing of visual input, facilitating a biologically inspired sensing paradigm that better reflects the continuous nature of real-world dynamics.

To evaluate the impact of this architectural difference, we formulated a trajectory prediction task[36,37] in which an artificial neural network (ANN) was trained to predict the coordinates of a moving circular object three steps into the future. Using experimental measurements, we constructed a mathematical model of the non-volatile memory behavior of the photonic-memory sensor and used it to generate synthetic sensor data. To test the experimental transferability of this numerical imaging model, we additionally validated it using experimentally reconstructed motion patterns based on measured ZnO-based synaptic-pixel responses. We then compared the photonic-memory sensor with a conventional image sensor within the same prediction framework by varying both the noise level and the number of input frames supplied to the ANN. To ensure that prediction performance was governed primarily by object motion, we considered a deliberately simplified visual scene consisting of a single moving object on a uniform background. For a fair comparison, the model was trained and evaluated separately on sensor data generated by the conventional image sensor and the photonic-memory sensor. For a comprehensive comparison, we constructed multiple datasets in which key parameters, such as the number of input frames and the noise level, were systematically varied. The noise level was controlled by applying Gaussian noise to the dataset images, simulating real-world sensor noise conditions. Each dataset consisted of frame sequences capturing object motion, with spatiotemporal features extracted based on the operational characteristics of each sensor type. To process these data, we optimized two separate models, each combining convolutional neural networks (CNNs) for feature extraction with long short-term memory (LSTM) networks for sequential data processing. Although the visual scene was deliberately kept simple, the prediction pipeline follows modern deep-learning-based trajectory prediction methods that are widely employed in real-world applications. Model performance was assessed on corresponding test datasets, using Root Mean Square Error (RMSE)[38] as the primary evaluation metric. Taken together, our findings lead to two key conclusions within this experimentally grounded numerical framework. First, the photonic-memory sensor enables reliable trajectory prediction from a single input frame, providing a clear advantage over the conventional sensor in this regime. Second, as the number of input frames increases, this advantage is not only lost but eventually reversed: the conventional image sensor ultimately achieves higher accuracy and more robust predictive performance. These results indicate that photonic memory alone is insufficient for optimal multi-frame trajectory prediction and must be complemented by additional device-level and architectural mechanisms.

II. MATERIALS AND METHODS

A. Optoelectronic artificial synapses

Optoelectronic artificial synapses are designed to replicate key properties of biological synapses, such as synaptic plasticity[39] and signal-processing capabilities. By mimicking these functions, they play a crucial role in neuromorphic computing systems, which aim to process information in a brain-like manner. Among the various materials used for their fabrication, nanocrystalline metal oxides have attracted significant attention because of their stability, scalability, and responsiveness to external stimuli.[4,40–43] These materials have also been successfully integrated into neuromorphic sensory systems, in which each pixel of an optoelectronic sensor operates as an artificial synapse, enabling direct optical signal processing in a biologically inspired fashion.[18,44–46] The properties of optoelectronic artificial synapses based on ZnO are governed by oxygen adsorption and desorption processes, as well as charge-carrier recombination and trapping, which regulate the conductivity dynamics under varying illumination conditions. These mechanisms enable ZnO-based structures to function as artificial synapses, exhibiting properties such as spike-like responses and the retention of information about previous stimuli. A detailed description of these physical processes and their role in synaptic behavior is provided in Ref. 47. Previous studies have shown that optoelectronic artificial synapses can be engineered to operate over a broad spectral range.[11,48] Although these systems often exhibit comparable ratios between conductance decay time and illumination duration, the absolute temporal scales can differ by several orders of magnitude.[48] This diversity in spectral sensitivity and temporal response provides additional flexibility for tailoring synaptic behavior to the specific requirements of different neuromorphic vision and optoelectronic information-processing applications.

B. Photonic memory sensor model based on experimental data

Based on these experimental data, we developed a numerical model of the artificial synapse conductance that reproduces its time evolution for variable illumination durations and temporal illumination protocols within the calibrated experimental condition. The connection between the experimentally measured single synaptic device and the numerical photonic-memory sensor model is schematically shown in Fig. 2. First, the conductance dynamics of a single ZnO-based optoelectronic synaptic device were experimentally characterized under controlled optical illumination and electrical readout conditions [Fig. 2(a)]. The measured potentiation and relaxation dynamics of this single structure were then used as the pixel-level response function in the numerical sensor model. In this model, each pixel of the photonic-memory sensor corresponds to an individual ZnO-based optoelectronic synaptic device [Fig. 2(b)]. Thus, the two-dimensional sensor array was constructed by assigning the experimentally measured illumination-history-dependent conductance dynamics to every pixel.

For all measurements, we used a previously established experimental setup[47] [Fig. 2(a)]. A ZnO-based optoelectronic artificial synapse was fabricated on an Al_2O_3 substrate with patterned platinum electrodes forming the active channel between the contacts. The device was connected in series with a constant reference resistor, and a constant bias voltage was applied across the series circuit. The voltage drop across the reference resistor was monitored by an analog-to-digital converter (ADC) interfaced with a personal computer. The current through the ZnO channel was then obtained from Ohm's law, and the channel conductance was calculated from the known device geometry, including electrode spacing and active area. In the key experiment, the ZnO structure was continuously illuminated with a 405 nm LED until its conductance reached saturation, after which the illumination was switched off and the subsequent dark-phase relaxation was recorded. The LED was positioned normal to the sample surface, and both the LED-sample distance and the driving current were kept constant to ensure uniform and reproducible illumination of the active region. During both the potentiation (under illumination) and decay (in darkness) phases, the conductance was recorded at a sampling frequency of 1 Hz, which was sufficient to resolve the relatively fast increase in conductance (on the order of 350 s) as well as the much slower relaxation (on the order of 14 000 s). As a result, two experimental curves

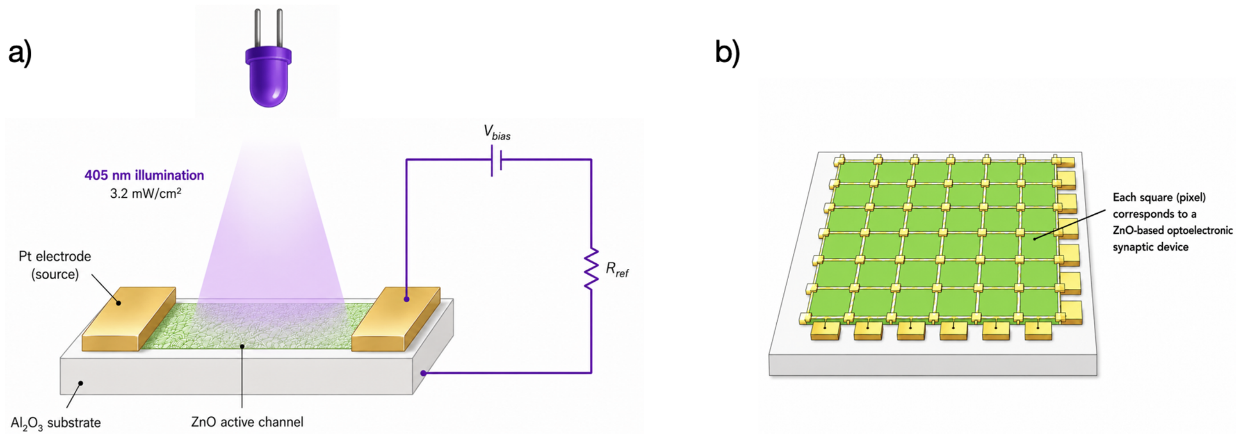

FIG. 2. Schematic representation of the experimental single-device characterization and its connection to the numerical photonic-memory sensor model. (a) Single ZnO-based optoelectronic synaptic device used for experimental conductance characterization. The device consists of a ZnO active channel on an Al_2O_3 substrate contacted by Pt electrodes. The active channel is illuminated by 405 nm light, while the conductance is electrically read out using a constant bias voltage and a reference resistor. (b) Numerical photonic-memory sensor model represented as a two-dimensional array of ZnO-based synaptic pixels. Each pixel corresponds to an individual optoelectronic synaptic device whose conductance follows the experimentally measured potentiation and relaxation dynamics.

were obtained, representing the evolution of the structure conductance during potentiation and decay. The measured conductance values were subsequently normalized to a scale from 0 to 255, corresponding to the full range of pixel brightness in an 8-bit image. This normalization allows the synaptic conductance to be directly interpreted as grayscale levels in neuromorphic image-processing tasks while preserving the relative dynamics of the potentiation and decay processes. During numerical image generation, the normalized pixel intensity was additionally clipped to the 8-bit image range,

$$I(t) = \min\left[255, \max\left(0, \sigma(t)\right)\right],$$

where $\sigma(t)$ is the normalized conductance response obtained from the fitted potentiation or relaxation function. Therefore, the polynomial potentiation approximation was not used as an unbounded long-time conductivity law; it served as an empirical interpolation of the measured potentiation dynamics before conversion to a bounded image intensity.

The potentiation and decay curves (Figs. 3 and 4) were approximated using simple analytical functions, chosen in line with previously reported phenomenological models of optoelectronic and photoresponsive structures[49–51] and optimized to reproduce the measured conductivity dynamics over the experimental time window. The potentiation process was described by the following third-degree polynomial function:[49]

$$\sigma_{\mathrm{pot}}(t) = A_{\mathrm{pot}} + B_{\mathrm{pot}}t + C_{\mathrm{pot}}t^2 + D_{\mathrm{pot}}t^3, \qquad (1)$$

where A_{pot}, B_{pot}, C_{pot}, and D_{pot} are adjustable parameters, and t is the time. The cubic polynomial was selected empirically because it provided the most accurate and numerically stable fit to the measured potentiation curve within the experimental time window. A saturating exponential function was also tested and is physically more appropriate for describing long-time saturation. However, within the calibrated interval considered here, the polynomial approximation reproduced the experimental data more accurately and was therefore retained. Since the simulated pixel intensity was clipped to the 8-bit range, the polynomial was used only as a bounded empirical interpolation of the measured response, not as an unconstrained long-time conductivity model.

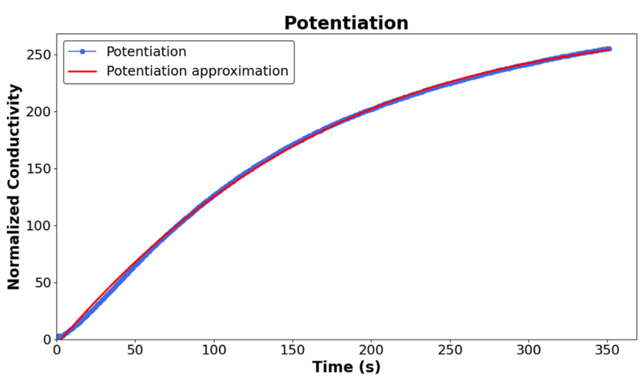

FIG. 3. Potentiation of normalized conductivity as a function of time and its analytical approximation.

The decay behavior of photoresponsive structures with similar dynamic characteristics is commonly modeled phenomenologically using a sum of exponential functions to account for several relaxation channels with different characteristic time scales.[50,51] Following this approach, we initially used a double-exponential function of the form

$$\sigma_{\mathrm{rel}}(t) = A_{\mathrm{rel},1}\,\exp\left(-\frac{t}{\tau_1}\right) + A_{\mathrm{rel},2}\,\exp\left(-\frac{t}{\tau_2}\right), \qquad (2)$$

where $A_{\mathrm{rel},1}$ and $A_{\mathrm{rel},2}$ are the amplitudes of the corresponding relaxation components, and τ_1 and τ_2 are their effective decay constants. In general, such a representation can capture contributions from multiple relaxation processes operating on different time scales.

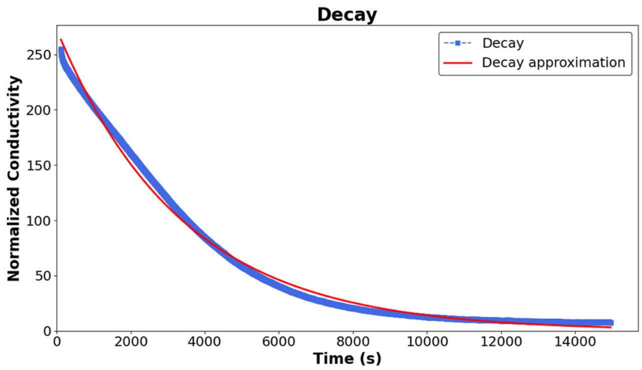

FIG. 4. Decay of normalized conductivity as a function of time and its analytical approximation.

For our ZnO-based structure, however, the fitting procedure yielded nearly identical values of τ_1 and τ_2, indicating that the experimental decay can be accurately described by a single characteristic time scale. As a result, Eq. (2) effectively reduces to a single-exponential dependence,

$$\sigma_{\mathrm{rel}}(t) = A_{\mathrm{rel}} \exp\left(-\frac{t}{\tau_{\mathrm{rel}}}\right), \tag{3}$$

where A_{rel} is the amplitude of the photoinduced excess conductivity at the beginning of the dark phase, and τ_{rel} is the effective relaxation time. In this phenomenological description, τ_{rel} summarizes the dominant recombination and de-trapping processes in the ZnO structure without attempting to resolve individual microscopic channels.

Table I presents the values of the adjustable parameters for the potentiation and decay processes, respectively.

We emphasize that both the polynomial potentiation model and the exponential decay model are phenomenological. They are not intended to reproduce the full physics of charge transport and trapping in ZnO, but rather to provide a compact and numerically efficient description of the experimentally observed conductivity dynamics. The fitting parameters reported here are specific to the calibrated illumination condition used in the experiment. The same modeling framework can be applied to other wavelengths or optical power densities, but the potentiation and relaxation parameters must then be recalibrated from the corresponding experimental conductance curves. This level of description is sufficient for our purpose, namely, to construct a numerical artificial synapse model that reproduces the device conductivity response for variable illumination durations and temporal illumination protocols within the calibrated regime.

TABLE I. Fitting coefficients for the potentiation and decay functions.

Process	Parameter	Value
Potentiation	A_{pot}	−4.91
	B_{pot}	1.62
	C_{pot}	-3.47×10^{-3}
	D_{pot}	2.72×10^{-6}
Decay	A_{rel}	272.79
	τ_{rel}	2.96×10^{3}

C. Experimental validation of the numerical photonic-memory model

To test whether the numerical photonic-memory imaging model produces experimentally transferable motion-encoding patterns, we performed an additional validation experiment based on experimentally reconstructed 5×5 response patterns. This validation was not intended to reproduce the full trajectory-prediction benchmark or to demonstrate a fully integrated camera-like sensor array. Such a demonstration would require a larger homogeneous ZnO sensor array with a sufficient number of independently characterized pixels. Instead, the experiment was designed to verify whether the motion-encoding mechanism predicted by the numerical model can also be reproduced from experimentally measured ZnO-based synaptic-pixel responses.

The ZnO-based optoelectronic synaptic pixels were characterized using the same conductance-readout principle as described in Sec. II B. The experimentally measured conductance was used as the physical observable and was subsequently converted into grayscale intensity values, consistent with the normalization procedure used in the numerical sensor model.

First, we generated a synthetic dataset corresponding to the motion of a point-like object across a photonic-memory sensor. The dataset was produced using the potentiation and relaxation dynamics introduced in Sec. II B. A two-layer multilayer perceptron (MLP) was trained on these model-generated grayscale patterns to recognize the direction of motion. The validation task was formulated as a four-class classification problem, corresponding to upward, downward, leftward, and rightward motion. The output layer contained four neurons, one for each direction.

Experimental point-like motion was then reconstructed by assigning a temporal illumination history to each pixel of a 5×5 grid. For a given direction of motion, each selected pixel was kept in darkness before the point-like object reached its position, illuminated for a fixed time while the object occupied that position, and then left in darkness to relax after the object moved away. In this way, the illumination history assigned to the pixels reproduced the stepwise motion of a point-like object across the reconstructed grid. Three illumination times per pixel were tested: 50, 120, and 150 s. These temporal regimes emulate different effective velocities of point-like motion. The resulting conductance values were converted into grayscale intensity values to form experimentally reconstructed motion-encoding patterns. A representative reconstructed ZnO-based response pattern is shown in Fig. 5.

The trained MLP was then evaluated on reconstructed ZnO-based motion-encoding patterns. The network was trained only on synthetic data generated by the numerical model; no reconstructed experimental patterns were used during training. The validation set included 12 patterns: four motion directions and three illumination times. All 12 patterns were classified correctly, indicating that the numerical model captures the main direction-dependent features reproduced by the measured ZnO-based synaptic-pixel dynamics.

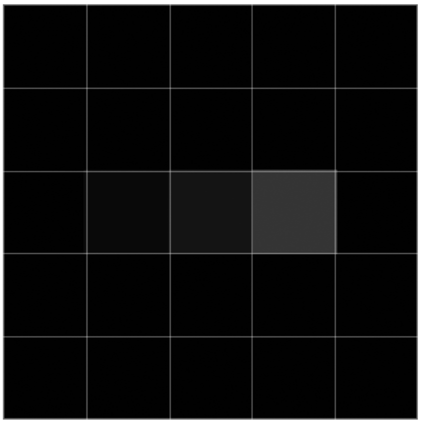

FIG. 5. Representative experimentally reconstructed grayscale response pattern based on ZnO-based optoelectronic synaptic-pixel dynamics under a temporal illumination protocol emulating point-like motion to the right. The grid lines indicate the reconstructed 5×5 pixel positions.

This validation confirms the motion-encoding mechanism used in the numerical photonic-memory imaging model, but does not constitute full experimental validation of the trajectory-prediction benchmark or a demonstration of a fully integrated sensor array. Therefore, the trajectory-prediction results below should be interpreted as proof-of-principle numerical results grounded in experimentally measured ZnO photonic-memory dynamics.

D. Trajectory prediction datasets

To compare the performance of conventional image sensors and photonic-memory sensors in the trajectory prediction task, it is first necessary to simulate how each sensor responds to object motion. These simulated data then serve as input for training and evaluating trajectory prediction models for both sensor types.

To this end, we simulated the motion of an object across a dark, spatially uniform background. The object moved at a constant speed and changed direction after traveling a random distance between 5 and 40 pixels.

Each segment of motion followed a linear trajectory defined by the equation,

$$y = kx + b, \tag{4}$$

where k is the trajectory slope and b is the initial offset. New values for k and b were selected from a predefined set (Table II) each time the direction changed.

Each generated frame was a 600×600 binary image with the object represented as a white circle. The object moved in discrete steps of one pixel, and each frame filename encoded the coordinates of the object's center in the next time step. The object was assumed to be positioned directly in front of the sensor matrix, illuminating only the pixels it covered. The visual scene intentionally contained no background texture or additional objects, so that trajectory prediction performance would be determined primarily by the motion of the target. This deliberately simplified setup allowed us to isolate the contribution of the photonic memory effect from confounding factors related to scene complexity.

TABLE II. Parameters for linear trajectories. Predefined values of slope k and offset b used in the trajectory generation algorithm.

Trajectory slope k	Initial offset b
1.0	0.0
2.0	1.0
−1.0	0.0
0.5	−2.0
3.0	5.0

The piecewise-linear trajectory generator was chosen as a controlled intermediate setting. The trajectories are not fixed deterministic straight lines, since the direction changes after randomly selected travel distances. At the same time, the short-horizon prediction problem remains well posed, which is important for isolating the effect of the sensor temporal response. More complex curvilinear trajectories can introduce additional confounding factors: deterministic curves may allow the network to rely on trajectory-family priors, whereas randomly varying curvature can make future motion weakly constrained by the available input frames. For this reason, the present benchmark focuses on controlled stochastic piecewise-linear motion.

The response of both sensor types was then modeled. The conventional sensor captured a sequence of discrete frames with no memory effect. In contrast, the photonic-memory sensor implemented a different mechanism. Each pixel exhibited a time-dependent response to incident illumination, governed by potentiation and relaxation functions.

As shown in Fig. 6, when the object was positioned directly in front of a concrete pixel, it was illuminated according to the potentiation function described by (1). Once the object moved away, the pixel underwent a decay process, following the behavior defined in (3). The photonic-memory sensor generated a smooth, continuous response to changes in object position rather than discrete measurements. This continuous output allows the sensor to encode the temporal context of object motion, unlike conventional sensors that rely on discrete frames. As illustrated in Fig. 7, the sensor's response retains information about the object's previous positions, thereby preserving motion history within the sensor output. This capability demonstrates the potential of photonic memory to bridge the gap between artificial sensing and biological perception.

To simulate realistic conditions in which sensor data are acquired at discrete time intervals, we introduced a temporal discretization. For both sensor types, we selected every nth frame from the full response sequence, where n denotes the discretization factor (i.e., only one out of every n consecutive frames is retained).

As a result, two datasets were formed:

- **Frame-Based Dataset**—consists of discrete image frames representing conventional sensor output.
- **Memory-Based Dataset**—contains temporally smoothed responses of the photonic-memory sensor.

We also investigated the impact of measurement noise on prediction performance. To this end, we added Gaussian noise to the

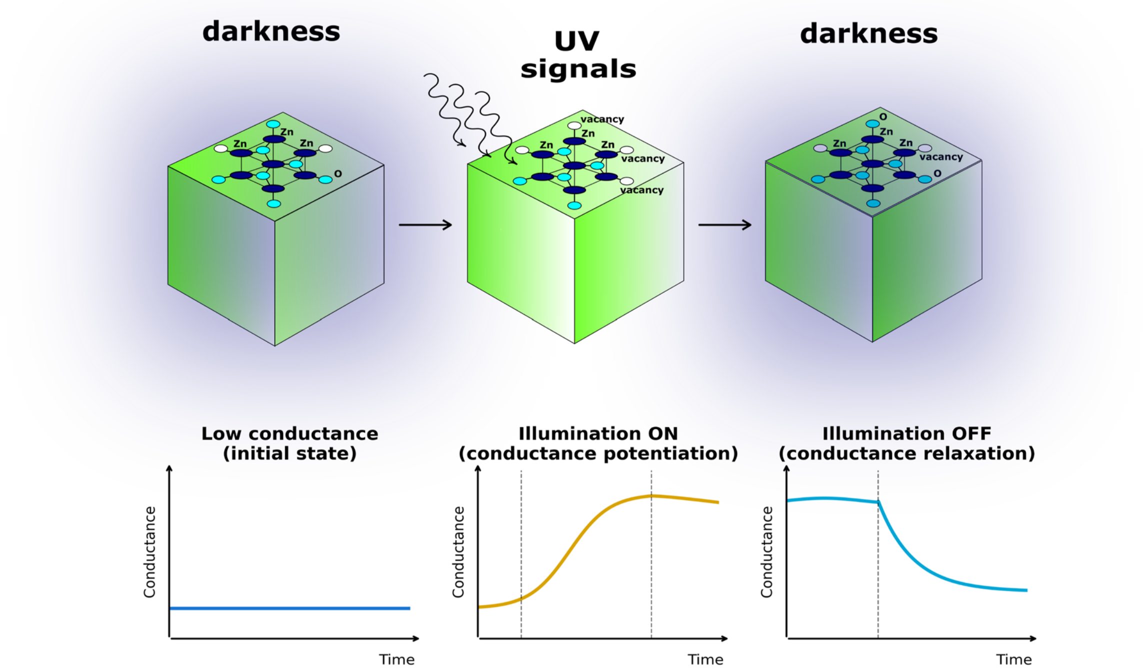

FIG. 6. Schematic illustration of the photonic-memory response of a single ZnO-based optoelectronic synaptic pixel during the initial dark state, illumination-induced potentiation, and post-illumination relaxation.

preformed datasets, using two configurations of mean and variance: either (80, 20) or (50, 10), respectively. In all cases, the underlying visual scene remained deliberately simple and texture-free, ensuring that the observed differences in prediction accuracy primarily reflect the presence or absence of photonic memory rather than variations in background structure.

E. Trajectory prediction model

The conventional sensor and the photonic-memory sensor differ only in the structure of the 2D image sensor array, while their processing and memory components remain identical. To ensure a fair comparison, we used two ANN models—corresponding to the conventional sensor and the photonic-memory sensor, respectively—with identical architecture. For each model, we selected an optimal set of hyperparameters based on performance on its respective dataset: a frame-based dataset for the conventional sensor and a memory-based dataset for the photonic-memory sensor. Once optimized, we evaluated both models on their respective test sets and compared their prediction accuracy metrics to assess how differences in sensor data influenced trajectory prediction performance.

Predictions were evaluated using various numbers of input frames, ranging from one to six. The model's objective was to predict the object's trajectory for the next three coordinates on the basis of the preceding frames.

To assess prediction accuracy, we used the RMSE metric, which quantifies the root of the mean squared Euclidean distance between predicted and ground-truth positions across all predicted frames,

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[(\hat{y}_i - y_i)^2\right]}, \tag{5}$$

where $\hat{y}_i$ are the predicted coordinates, and y_i are the ground-truth coordinates.

The model follows an end-to-end approach, where both feature extraction and coordinate prediction are performed within a single framework. The input data belong to the space $\mathscr{X}_n = (\mathbb{R}^{H\times W})^n$, where n denotes the number of consecutive input frames that contain a moving circle. Each input sequence is represented as $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in \mathscr{X}_n$, where each frame $z_t \in \mathbb{R}^{H\times W}$ represents a single frame as a pixel matrix of size $H \times W$, as shown in Fig. 7.

The model predicts the coordinates of the circle in the future frames, which belong to the two-dimensional Euclidean space $\mathbb{R}^2$. It functions as an operator $A : \mathscr{X}_n \to \mathbb{R}^6$, generalizing to different input sequence lengths.

To realize this operator, we implemented a neural network that processes the input sequence in three stages. First, each frame is passed through two convolutional layers to extract informative spatial features; this approach leverages the proven efficiency of convolutional neural networks in image feature extraction.[52] The resulting per-frame features are then input sequentially into a recurrent neural network module. They are widely used for modeling sequential data, as they can extract features and capture long-term

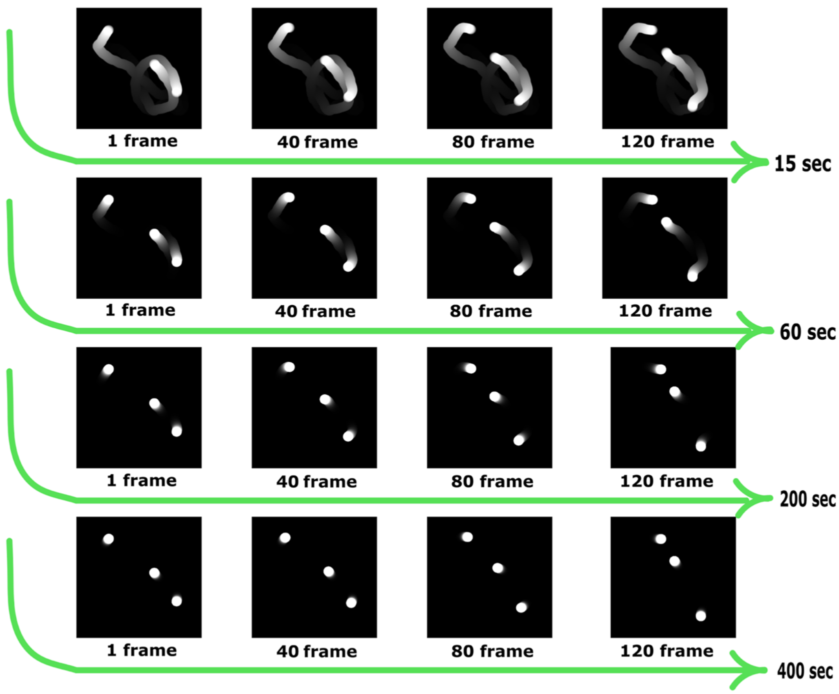

FIG. 7. Images generated from numerical experiments illustrating the response of the photonic-memory sensor at different temporal parameters (object movement speeds). "sec." refers to seconds.

dependencies in time series. However, their ability to learn from extended sequences is often hindered by the vanishing gradient problem, which affects the stability and efficiency of training.[53] To address this limitation, we adopted an LSTM network, which enables long-range dependency modeling and improved training stability.[54] The output of the LSTM is finally mapped by a linear layer to yield the predicted coordinates for future frames. Unlike sequence-to-sequence approaches that use both an LSTM encoder and decoder,[55] our model predicts a fixed-size vector of future positions, so only a single LSTM encoding layer is required.

The model was trained using the mean squared error loss function, which quantifies the average squared difference between the predicted and ground-truth coordinates,

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2, \tag{6}$$

where $\hat{y}_i$ are the predicted coordinates, and y_i are the ground-truth coordinates.

To ensure optimal performance, key hyperparameters—including learning rate, batch size, number of convolutional filters, convolutional kernel size, stride, padding, LSTM layer count, LSTM hidden-state dimensionality, and dropout—were optimized via Bayesian optimization. The following search ranges were used: batch size $B \in \{8, 16, 32, 64\}$, number of output channels in the first and second convolutional layers $N_{f1} \in [2, 16]$ and $N_{f2} \in [2, 32]$, convolutional kernel size $K \in \{3, 5, 7\}$, stride $S \in \{1, 2\}$, padding $P \in \{0, 1, 2, 3\}$, LSTM hidden-state dimensionality $H \in [32, 256]$, number of LSTM layers $L \in [1, 4]$, dropout $D \in [0, 0.5]$, and learning rate $\eta \in [10^{-5}, 10^{-2}]$. For each trial, the model was trained on the training set and evaluated on the validation set. The neural network was trained using the AdamW optimizer and the MSE loss function, while the validation objective for hyperparameter optimization was the absolute trajectory error (ATE), defined as the mean Euclidean distance between the predicted and ground-truth trajectories,

$$\mathrm{ATE} = \frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{p}_i^{\mathrm{pred}} - \mathbf{p}_i^{\mathrm{gt}}\right\|_2, \tag{7}$$

where $\mathbf{p}_i^{\mathrm{pred}}$ and $\mathbf{p}_i^{\mathrm{gt}}$ are the predicted and ground-truth object positions at the ith predicted time step, respectively, and N is the number of predicted time steps.

We systematically evaluated the predictive performance of two independently trained models: one using data from the conventional sensor, and the other using data from the photonic-memory sensor. The experimental design included all possible combinations of two discretization levels ($n = 5$ and $n = 20$), two noise configurations (50, 10) and (80, 20), and the noise-free scenario. For the photonic-memory sensor, we further considered three different object-motion regimes (15, 60, and 400 s/pixel), corresponding to the illumination time per pixel before the object shifted to the next pixel. In contrast, for the conventional sensor, this parameter was not varied because the conventional memoryless sensor produces instantaneous image snapshots and has no conductance state that accumulates or relaxes during illumination. For each unique scenario, we trained and tested models on data from each sensor type separately, independently varying the number of input frames provided to the model from one to six.

It should be noted that, in the photonic-memory simulations, the varied object-motion regime determines the illumination time per pixel and therefore controls the strength and spatial extent of the residual memory trace. This parameter affects the generated photonic-memory images because it changes how much residual conductance remains from previous object positions. In contrast, the conventional sensor is modeled as a memoryless image sensor: each frame represents an instantaneous snapshot of the object position, and changing the illumination time per pixel does not introduce an additional device-level state. Therefore, this parameter affects the generated photonic-memory images, but not the conventional images.

The LSTM implementation was identical for all sensor types and illumination-time regimes. For a given number of input frames, the model received a sequence of n images, with $n = 1, \ldots, 6$. Each frame was first processed by the convolutional feature-extraction block, and the resulting sequence of feature vectors was then supplied to the LSTM. Thus, the LSTM sequence length was determined only by the number of input frames, while the illumination time per pixel affected only the intensity distribution within the photonic-memory frames.

For every combination of scenario, sensor type, and input frame count, we performed separate hyperparameter optimization using the same model class, training procedure, validation objective, and optimization budget of 100 Bayesian-optimization iterations. This protocol was chosen to compare the best attainable trajectory-prediction performance of each sensor representation under a controlled applied-ML setting, rather than to compare under-optimized neural networks. In total, this comprehensive experimental design encompassed 144 unique conditions.

III. RESULTS

A. Trajectory prediction performance

To compare the performance of the two sensor types, we computed, for each number of input frames, the difference in RMSE between the corresponding models under every experimental scenario. The results are summarized in Fig. 8(b), which shows, for each frame count, the percentage of cases in which the model trained and tested on data from the sensor with photonic memory achieved a lower RMSE than the model based on the conventional sensor.

Even when only a single frame from the conventional sensor is available, the model achieves non-trivial trajectory prediction accuracy. This is a consequence of the highly constrained nature of the synthetic scene: a single circular object moves according to simple, smooth motion rules within a finite field of view. Under these conditions, the instantaneous position of the object already carries statistical information about its likely past and future motion. For example, proximity to a boundary or corner restricts the admissible directions and magnitudes of subsequent displacements. During training, the network learns these regularities and effectively approximates the conditional expectation of future positions given the current spatial configuration, which enables approximate trajectory prediction from spatial cues alone, even in the absence of explicit temporal information.

As illustrated in Fig. 8(b), the sensor with photonic memory exhibits a pronounced advantage when only a single frame is available: in more than 85% of experimental configurations, it yields a lower prediction error than the conventional sensor. This substantial margin highlights the value of the memory effect, which allows the sensor to exploit residual information from prior exposures, thereby compensating for the lack of explicit temporal context.

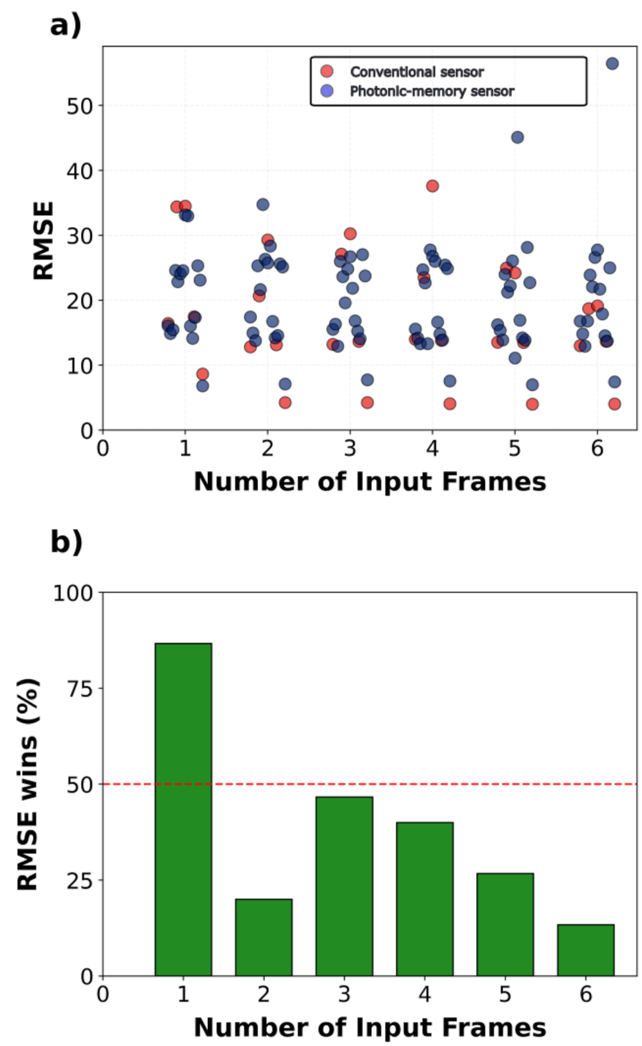

FIG. 8. (a) RMSE as a function of the number of input frames. Each point corresponds to a unique combination of discretization level (n = 5 or 20), noise level (50, 10; 80, 20; or noise-free), and, for the photonic-memory sensor, object speed (15, 60, or 400 steps/pixel). Red markers indicate the conventional sensor (speed fixed), while blue markers represent the sensor with photonic memory. (b) Percentage of cases in which the sensor with photonic memory achieves a lower RMSE than the conventional sensor, as a function of the number of input frames.

As the number of input frames increases, however, this advantage is substantially reduced. While the sensor with photonic memory is clearly superior in the single-frame case, its success rate decreases for larger numbers of frames and falls below 50%. In other words, once temporal information is explicitly provided through a sequence of frames, the conventional sensor performs comparably or even better, and photonic memory no longer offers a systematic benefit for trajectory prediction.

The pronounced drop at two input frames in Fig. 8(b) reflects a specific property of our trajectory-prediction task rather than statistical noise. For the conventional sensor, two frames already provide two clean positions, which almost uniquely determine the local velocity vector and thus enable a nearly optimal short-horizon extrapolation. In contrast, for the photonic-memory sensor, this is the most unfavorable configuration: when the object changes direction between the two frames, the temporal integration in the photonic synapse merges the pre- and post-turn segments into a single blurred trace that is poorly matched by a simple linear motion model. At longer sequence lengths, additional frames partially resolve this ambiguity and the relative performance of the photonic-memory sensor recovers. Therefore, the non-monotonic behavior at two frames is consistent with the interplay between the engineered temporal kernel of the photonic sensor and the piecewise-linear structure of the trajectories, and does not contradict the overall conclusion that photonic memory provides a clear benefit only in the single-frame regime.

IV. DISCUSSION

From a sensing perspective, the photonic-memory sensor effectively shifts part of the motion-encoding process from the digital domain to the optical front end: residual charges from preceding exposures encode a short history of motion directly in the spatial intensity pattern. When coupled with a sufficiently expressive neural model, this representation allows trajectory-relevant temporal information to be extracted from a single spatial frame, rather than solely from explicit frame-to-frame differences. Such a mechanism is particularly attractive in regimes where frame rate or bandwidth are severely constrained, or where latency requirements limit the number of frames that can be acquired before a decision must be made.

At the same time, our results indicate that for trajectory prediction from multi-frame image sequences, simple photonic memory alone is not sufficient to systematically outperform a conventional sensor. Once several clean, temporally separated frames are available, the benefits of optical integration diminish, and the superposition of trajectories can even hinder inference. This suggests that, for motion-related tasks, neuromorphic optoelectronic sensors should combine photonic memory with additional biologically inspired properties—most notably synaptic-like depression and related adaptive mechanisms—so that the effective temporal response can be tuned to the task. In this sense, our study points toward the development of device architectures with a richer and more flexible set of dynamical properties, accompanied by more diverse and realistic motion datasets, as a promising direction for advancing neuromorphic vision sensors beyond the capabilities of conventional passive imagers.

Several limitations of the present study should be noted. First, the trajectory-prediction benchmark was evaluated in a numerical imaging framework rather than on a fully integrated large-scale sensor array. This choice reflects the main objective of the work: to isolate and quantify how photonic memory, as a device-level temporal response, affects trajectory prediction under controlled conditions. The numerical model is grounded in experimentally measured ZnO photonic-memory dynamics and is supported by an additional validation of the underlying motion-encoding mechanism using experimentally reconstructed ZnO-based response patterns. However, this validation should not be interpreted as a full experimental demonstration of the complete trajectory-prediction pipeline.

Second, the absolute timescale of the ZnO photonic-memory response should be interpreted in relation to the purpose of the study. The present work does not aim to demonstrate a ready-to-use high-speed vision device. Rather, it uses experimentally measured ZnO dynamics as a calibrated memory kernel to evaluate how sensor-level temporal integration can influence trajectory prediction. Faster optoelectronic synaptic responses have been reported in other material and device platforms,[56] suggesting that the same modeling principle could be implemented at shorter timescales. In such implementations, the functional role of photonic memory would remain the same, while the absolute temporal scale of the potentiation and relaxation dynamics would be matched to the characteristic timescale of image acquisition and object motion.

Third, the benchmark intentionally uses a simplified motion-prediction setting. A single moving object on a uniform background was chosen to isolate the contribution of photonic memory without additional confounding factors from complex scenes. More realistic scenarios, such as textured backgrounds, occlusions, multi-object motion, or strongly nonlinear trajectories, may require additional sensor-level mechanisms and more specialized prediction pipelines. Extending the analysis to such regimes remains an important direction for future work.

V. CONCLUSIONS

In this work, we compared trajectory-prediction performance using image sequences from a conventional sensor and from a ZnO-based photonic-memory sensor in an experimentally grounded numerical framework. Across a broad set of conditions (noise levels, discretization parameters, and object speeds), we found that photonic memory provides a clear advantage in the single-frame regime: when only one frame is available, the photonic-memory sensor systematically yields lower prediction error than the conventional sensor in the majority of tested configurations (exceeding 85%).

However, this advantage does not extend to multi-frame inputs. When two or more frames are provided, the photonic-memory sensor is outperformed by the conventional sensor in most tested configurations, indicating that explicit temporal sampling of clean, non-overlapping frames is more beneficial for sequence-based trajectory prediction than optical integration of motion traces. Taken together, these results show that, in the context of the numerical trajectory-prediction task considered here, photonic memory mainly benefits single-frame inference, whereas in the multi-frame regime conventional sensing with explicit temporal sampling generally provides superior performance.

SUPPLEMENTARY MATERIAL

The supplementary material contains a speed-resolved paired RMSE analysis comparing the photonic-memory sensor with the conventional image sensor at different illumination times and numbers of input frames.

ACKNOWLEDGMENTS

The study was supported by the grant of the Russian Science Foundation Grant No. 25-12-00355, https://rscf.ru/project/25-12-00355/. Authors also thank the Noncommercial Foundation for support of Science and Education "INTELLECT."

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Fedor L. Ivanov: Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Writing – original draft (equal). **Viktor V. Krasnikov**: Conceptualization (equal); Investigation (equal); Writing – original draft (equal). **Andrey A. Grunin**: Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Supervision (equal). **Artem V. Chetvertukhin**: Conceptualization (equal); Formal analysis (equal); Project administration (equal); Resources (equal); Supervision (equal). **Andrey A. Fedyanin**: Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data and code that support the findings of this study are openly available in the GitHub repository "Optoelectronic Sensors with Photonic Memory for Trajectory Prediction: Data and Code" at https://github.com/tortolla/photonic-memory-trajectory. The repository contains raw ZnO calibration traces, trajectory-generation scripts, photonic-memory sensor simulation code, Gaussian-noise generation scripts, reference training notebooks, experimental-grid configuration files, CSV tables with trajectory-prediction metrics, and representative generated image frames illustrating the dataset format. The complete generated image-frame datasets are not deposited separately because they are produced directly by the released data-generation scripts from the provided configuration files and calibration data.

REFERENCES

[1]H.-H. Lee, J.-S. Ro, K.-N. Kim, H.-L. Park, and T.-W. Lee, Curr. Opin. Solid State Mater. Sci. **35**, 101215 (2025).
[2]C. Choi, G. J. Lee, S. Chang, Y. M. Song, and D.-H. Kim, ACS Nano **18**, 1241 (2024).
[3]D. A. Atchison, *Optics of the Human Eye* (CRC Press, 2023).
[4]J. Yu, Y. Wang, S. Qin, G. Gao, C. Xu, Z. Lin Wang, and Q. Sun, Mater. Today **60**, 158 (2022).
[5]A. Mehonic and J. Eshraghian, APL Mach. Learn. **1**, 020401 (2023).
[6]X. Zou, S. Xu, X. Chen, L. Yan, and Y. Han, Sci. China Inf. Sci. **64**, 160404 (2021).
[7]L. El Srouji, A. Krishnan, R. Ravichandran, Y. Lee, M. On, X. Xiao, and S. J. Ben Yoo, APL Photonics **7**, 051101 (2022).
[8]F. Liao, F. Zhou, and Y. Chai, J. Semicond. **42**, 013105 (2021).
[9]D. Ielmini and S. Ambrogio, Nanotechnology **31**, 092001 (2019).
[10]K. Aboumerhi, A. Güemes, H. Liu, F. Tenore, and R. Etienne-Cummings, J. Neural Eng. **20**, 041004 (2023).
[11]J. Zhang, P. Guo, Z. Guo, L. Li, T. Sun, D. Liu, L. Tian, G. Zu, L. Xiong, J. Zhang, and J. Huang, Adv. Funct. Mater. **33**, 2302885 (2023).
[12]H.-L. Park, Y. Lee, N. Kim, D.-G. Seo, G.-T. Go, and T.-W. Lee, Adv. Mater. **32**, 1903558 (2020).
[13]J. Zhu, T. Zhang, Y. Yang, and R. Huang, Appl. Phys. Rev. **7**, 011312 (2020).
[14]A. E. Pereda, Nat. Rev. Neurosci. **15**, 250 (2014).
[15]N. Ilyas, J. Wang, C. Li, D. Li, H. Fu, D. Gu, X. Jiang, F. Liu, Y. Jiang, and W. Li, Adv. Funct. Mater. **32**, 2110976 (2022).
[16]X. Wang, Y. Zong, D. Liu, J. Yang, and Z. Wei, Adv. Funct. Mater. **33**, 2213894 (2023).
[17]G. Cao, P. Meng, J. Chen, H. Liu, R. Bian, C. Zhu, F. Liu, and Z. Liu, Adv. Funct. Mater. **31**, 2005443 (2021).
[18]Q. Liu, L. Yin, C. Zhao, Z. Wu, J. Wang, X. Yu, Z. Wang, W. Wei, Y. Liu, I. Z. Mitrovic *et al.*, Nano Energy **97**, 107171 (2022).
[19]X. Shan, C. Zhao, Y. Lin, J. Liu, X. Zhang, Y. Tao, C. Wang, X. Zhao, Z. Wang, H. Xu, and Y. Liu, Appl. Phys. Lett. **121**, 263501 (2022).
[20]S.-H. Luo, C. Fu, C.-Y. Zhang, Z.-Y. Li, Z.-G. Zhu, J. Wang, Y. Wang, G. He, L.-B. Luo, and F.-X. Liang, IEEE Trans. Electron Devices **71**, 4745 (2024).
[21]Z. Jiang, Z. Luo, K. Liu, P. Jiao, W. Liu, and Y. Liu, IEEE Trans. Electron Devices **71**, 5456 (2024).
[22]Z. Li, R. Zhang, Z. Song, J. Sun, Z. Wang, D. Wu, and K. Wang, IEEE Trans. Electron Devices **70**, 5169 (2023).
[23]X. Liu, D. Li, Y. Wang, D. Yang, and X. Pi, APL Mach. Learn. **1**, 031501 (2023).
[24]J. Chen, H. Du, H. Qu, H. Gao, Y. Gu, Y. Zhu, W. Ye, J. Zou, H. Wang, and X. Zou, APL Mach. Learn. **2**, 026113 (2024).
[25]F. Zhou, Z. Zhou, J. Chen, T. H. Choy, J. Wang, N. Zhang, Z. Lin, S. Yu, J. Kang, H.-S. P. Wong, and Y. Chai, Nat. Nanotechnol. **14**, 776 (2019).
[26]F. Ferrarese Lupi, M. Rosero-Realpe, A. Ocarino, F. Frascella, G. Milano, and A. Angelini, Adv. Mater. **37**, 2501813 (2025).
[27]J. Chen, J. Zhu, P. Li, J. Hu, Q. Li, Z.-H. Wang, Z. Zhang, H. Li, S. Lin, X. Yue *et al.*, Nano Lett. **25**, 13909 (2025).
[28]S. H. Lee, H. J. Lee, D. Jeon, H.-J. Kim, and S.-N. Lee, Nanomaterials **14**, 1501 (2024).
[29]Q. Li, T. Wang, X. Hu, X. Wu, H. Zhu, L. Ji, Q. Sun, D. W. Zhang, and L. Chen, IEEE Electron Device Lett. **43**, 1089 (2022).
[30]L. Zhao, Z. Fan, S. Cheng, L. Hong, Y. Li, G. Tian, D. Chen, Z. Hou, M. Qin, M. Zeng *et al.*, Adv. Electron. Mater. **6**, 1900858 (2020).
[31]F. Ma, Y. Zhu, Z. Xu, Y. Liu, X. Zheng, S. Ju, Q. Li, Z. Ni, H. Hu, Y. Chai *et al.*, Adv. Funct. Mater. **30**, 1908901 (2020).
[32]J.-H. Oh, J.-H. Kim, H.-J. Kim, Y.-H. Kim, K.-K. Kim, and S.-N. Lee, Adv. Intell. Syst. **5**, 2300350 (2023).
[33]K. Chen, H. Hu, I. Song, H. B. Gobeze, W.-J. Lee, A. Abtahi, K. S. Schanze, and J. Mei, Nat. Photonics **17**, 629 (2023).
[34]J. Chen, Z. Zhou, B. J. Kim, Y. Zhou, Z. Wang, T. Wan, J. Yan, J. Kang, J.-H. Ahn, and Y. Chai, Nat. Nanotechnol. **18**, 882 (2023).
[35]M. He and X. Tang, Photonics **12**, 435 (2025).
[36]A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, and K. O. Arras, Int. J. Robot. Res. **39**, 895 (2020).
[37]Y. Huang, J. Du, Z. Yang, Z. Zhou, L. Zhang, and H. Chen, IEEE Trans. Intell. Veh. **7**, 652 (2022).
[38]Z. Zhang and D. Scaramuzza, in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2018), pp. 7244–7251.
[39]T.-H. Lee, H.-G. Hwang, J.-U. Woo, D.-H. Kim, T.-W. Kim, and S. Nahm, ACS Appl. Mater. Interfaces **10**, 25673 (2018).
[40]S. W. Cho, C. Jo, Y.-H. Kim, and S. K. Park, Nano-Micro Lett. **14**, 203 (2022).
[41]N. Kumar, M. Patel, T. T. Nguyen, P. Bhatnagar, and J. Kim, Mater. Today Chem. **23**, 100681 (2022).
[42]J. Zhou, N. Liu, L. Zhu, Y. Shi, and Q. Wan, "Energy-efficient artificial synapses based on flexible IGZO electric-double-layer transistors," IEEE Electron Device Lett. **36**(2), 198–200 (2015).
[43]R. Oommen, K. Jose, A. J G, S. Gandhi, V. Anand, and A. R. Nair, IEEE Sens. Lett. **8**, 1 (2024).

[44]Q. Liu, L. Yin, C. Zhao, J. Wang, Z. Wu, H. Lei, Y. Liu, B. Tian, Z. Zhang, Z. Zhao *et al.*, Nano Energy **102**, 107686 (2022).
[45]K. Liang, R. Wang, B. Huo, H. Ren, D. Li, Y. Wang, Y. Tang, Y. Chen, C. Song, F. Li *et al.*, ACS Nano **16**, 8651 (2022).
[46]P. S. Subin, P. S. Midhun, A. Antony, K. J. Saji, and M. K. Jayaraj, Mater. Today Commun. **33**, 104232 (2022).
[47]I. S. Balashov, A. A. Chezhegov, A. S. Chizhov, A. A. Grunin, K. V. Anokhin, and A. A. Fedyanin, Opto-Electron. Sci. **2**, 230016 (2023).
[48]X. Han, Z. Xu, W. Wu, X. Liu, P. Yan, and C. Pan, Small Struct. **1**, 2000029 (2020).
[49]P. Mougkogiannis, N. R. Kheirabadi, and A. Adamatzky, New J. Chem. **48**, 17650 (2024).
[50]P. Li, M. Zhang, Q. Zhou, Q. Zhang, D. Xie, G. Li, Z. Liu, Z. Wang, E. Guo, M. He *et al.*, Nat. Commun. **15**, 3257 (2024).
[51]M. M. Islam, D. Dev, A. Krishnaprasad, L. Tetard, and T. Roy, Sci. Rep. **10**, 21870 (2020).
[52]Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: Analysis, applications, and prospects," IEEE Trans. Neural Networks Learning Systems **33**(12), 6999–7019 (2022).
[53]H. Salehinejad, S. Sankar, J. Barfett, E. Colak, and S. Valaee, "Recent advances in recurrent neural networks," arXiv:1801.01078 [cs.NE] (2018).
[54]Y. Yu, X. Si, C. Hu, and J. Zhang, Neural Comput. **31**, 1235 (2019).
[55]I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Advances in Neural Information Processing Systems*, 27 (Curran Associates, Inc., 2014), pp. 3104–3112.
[56]H. Jang, S. Ju, S. Lee, J. Choi, U. Byun, K. Min, M. Rasheed, and S. Kim, Biomimetics **10**, 584 (2025).