



Supplementary Materials for **Spectral kernel machines with electrically tunable photodetectors**

Dehui Zhang *et al.*

Corresponding authors: Aydogan Ozcan, ozcan@ucla.edu; Ali Javey, ajavey@berkeley.edu

Science **390**, eady6571 (2025)
DOI: 10.1126/science.ady6571

The PDF file includes:

Supplementary Text
Figs. S1 to S19
Tables S1 to S3
References

Other Supplementary Material for this manuscript includes the following:

Movie S1

Supplementary Text

1. SKM Analysis

In this analysis, we aim to demonstrate that the positivity or negativity of the time-averaged photocurrent $\overline{I_{ph}}$ can serve as an indicator of the predicted class in a binary classification task using spectral information. We adopt a bottom-up approach: first presenting the optimal solution in the kernel machine model run in the digital processors (57), and then discussing the feasibility of implementing this design using our SKM.

Problem Formulation

Given a training dataset of N points, in the format of $\{S_i, z_i\}$, where each $S_i \in \mathbb{R}^{K \times 1}$ is a spectrum and $z_i \in \{-1, 1\}$ indicates the class of S_i , our objective is to find a hyperplane that separates the two classes. Specifically, the hyperplane satisfies:

$$\beta^T S_i - c = 0 \quad (\text{S1})$$

where $\beta \in \mathbb{R}^{K \times 1}$ is the normal vector to the hyperplane, and $c \in \mathbb{R}$ is a bias term. To maximize the margin, defined as $\frac{2}{\|\beta\|}$, we imposed the following constraints:

$$\beta^T S_i - c > 1, \text{ if } z_i = 1 \quad (\text{S2})$$

$$\beta^T S_i - c < -1, \text{ if } z_i = -1 \quad (\text{S3})$$

This equation can be rewritten as:

$$z_i(\beta^T S_i - c) > 1, \forall i \quad (\text{S4})$$

To allow for misclassification with training samples, we introduced slack variables $\alpha_i \geq 0$ and modify the constraint:

$$z_i(\beta^T S_i - c) > 1 - \alpha_i \quad (\text{S5})$$

The hinge loss function was then defined as:

$$\max(0, 1 - z_i(\beta^T S_i - c)) \quad (\text{S6})$$

The corresponding optimization problem can be written as:

$$\underset{\beta, c, \alpha}{\text{minimize}} \left(\frac{1}{2} \|\beta\|_2^2 + \sum_{i=1}^N \alpha_i \right)$$

$$s.t. \quad z_i(\beta^T S_i - c) \geq 1 - \alpha_i, \quad \alpha_i \geq 0, i = 1, 2, \dots, N \quad (\text{S7})$$

Dual Problem and Kernel Transformation

To solve the primal problem, we derived its dual formulation. The Lagrangian is expressed as:

$$L(\beta, c, \alpha, w, \xi) = \frac{1}{2} \|\beta\|_2^2 + \sum_{i=1}^N \alpha_i + \sum_{i=1}^N w_i (1 - \alpha_i - z_i(\beta^T S_i - c)) - \sum_{i=1}^N \xi_i \alpha_i \quad (\text{S8})$$

where $w_i \geq 0$ and ξ_i are Lagrange multipliers. By setting $\frac{\partial L}{\partial \beta} = 0$, $\frac{\partial L}{\partial c} = 0$, and $\frac{\partial L}{\partial \alpha_i} = 0$, we obtained (58):

$$\beta = \sum_{i=1}^N w_i z_i S_i \quad (\text{S9})$$

Substituting these equations into the Lagrangian yielded the dual problem:

$$\begin{aligned} & \underset{w}{\text{maximize}} \left(\sum_{i=1}^N w_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_i w_j z_i z_j S_i^T S_j \right) \\ & s.t. \quad \sum_{i=1}^N w_i z_i = 0, \quad 0 \leq w_i \leq 1, i = 1, 2, \dots, N \end{aligned} \quad (\text{S10})$$

This quadratic programming problem can be solved using standard convex optimization solvers. Once w_i values are determined, we substituted them back into Eq. S9 to compute β . For a new test spectrum $P \in \mathbb{R}^{K \times 1}$, the predicted class is:

$$z = \text{sign}(w \beta^T P - c) = \text{sign} \left(\left(\sum_{i=1}^N w_i z_i S_i^T \right) P - c \right) \quad (\text{S11})$$

Using a kernel function $k(P, S_i) = S_i^T P$ and omitting DC constant c , the decision rule simplified to:

$$z = \text{sign} \left(\sum_{i=1}^N w_i z_i k(P, S_i) \right) \quad (\text{S12}).$$

Relating to Time-Averaged Photocurrent

In the main text (Eq. 3), the time-averaged photocurrent was expressed as:

$$\overline{I_{ph}} = \int P(\lambda) \cdot Q(\lambda) \cdot d\lambda$$

In matrix form:

$$\overline{I_{ph}} = Q^T P \quad (\text{S13})$$

By analogy with the decision rule, if $Q = \sum_{i=1}^N w_i z_i S_i$, we can classify directly via:

$$\text{sign}(\overline{I_{ph}}) = \text{sign}(Q^T P) \quad (\text{S14})$$

Practical Implementation

In physical systems, the spectral encoding vector Q can be discretized along two dimensions: the temporal dimension (number of spectra used for spectral encoding in a time sequence, i.e., number of illuminations used) $j \in \{1, 2, \dots, J\}$ and spectral feature dimension (number of resolved wavelengths) $k \in \{1, 2, \dots, K\}$. Hence, we have:

$$Q_k = \sum_{j=1}^J F_{jk} R_{jk} \Gamma_j, k = 1, 2, \dots, K \quad (\text{S15})$$

where $F \in \mathbb{R}^{J \times K}$ is the encoding function, i.e., illumination; $R \in \mathbb{R}^{J \times K}$ is the trainable responsibility matrix of the sensor; $\Gamma \in \mathbb{R}^J$ is the readout bandwidth limited response function.

For most photodetectors, the responsivity spectra remain invariant in shape when tuned electrically, with changes manifesting only as a scaling factor controlled by V_{ctr} . Consequently, the responsivity can be expressed as

$$R_{jk} = \alpha_k r_j \quad (\text{S16})$$

where α_k is a factor determined solely by the light-absorbing material and is independent of V_{ctr} ; r_j represents the tunable factor controlled by V_{ctr} and associated with the j -th illumination.

Given this equation, the problem reduced to solving the following system of linear equations:

$$Ar = b \quad (\text{S17})$$

where $A = \begin{bmatrix} F_{11}\alpha_1\Gamma_1 & \dots & F_{J1}\alpha_1\Gamma_J \\ \vdots & \ddots & \vdots \\ F_{1K}\alpha_K\Gamma_1 & \dots & F_{JK}\alpha_K\Gamma_J \end{bmatrix} \in \mathbb{R}^{K \times J}$, $r \in \mathbb{R}^J$, and $b = \sum_{i=1}^N \lambda_i z_i S_i \in \mathbb{R}^K$. This system of K equations and J unknown variables $\{r_j\}$ has solutions depending on the relationship between b and the column space of A .

Feasibility of Solutions

If the b lies within the column space of A , the system would have at least one solution, allowing for the determination of r , which is related to the V_{ctr} . However, in the more general scenario where b does not lie in the column space of A , there is no exact solution. This situation arises because each column of A represents the spectrum of a spectral encoding scaled by $\alpha\Gamma$ and the number of encoding spectra (J) is limited. In contrast, b resides in a K -dimension space, which can be infinitely fine spectral structures. Therefore, J is typically much smaller than K , and the system becomes overdetermined, meaning it lacks an exact solution.

Approximate Solutions and Practical Considerations

Practically, we can still find an approximate solution by minimizing the residual subject to physical constraints, such as the allowable range of r_j . The optimization problem can be formulated as:

$$\begin{aligned} \min_r & \|Ar - b\|_2^2 \\ \text{s.t. } & r_{min} \leq r_j \leq r_{max}, j = 1, 2, \dots, J \end{aligned} \quad (\text{S18})$$

This objective function allows us to determine r and derive the corresponding V_{ctr} . Therefore, the entire process can be divided into two optimization problems. We acknowledge that the optimal solution for Eq. S7 (the first optimization problem solving for Q) may not align perfectly with the constraints imposed in the Eq. S18 (the second optimization problem solving for r). In other words, there may be a feasible solution for the second problem that performs better than the approximated solution solved from Eq. S18. Nonetheless, the above analysis provides an upper-bound performance for our physical system. This upper bound is particularly significant as it establishes

a benchmark for evaluating the physical feasibility of the SKM. By understanding the theoretical limits of our system, we can gauge how closely practical implementations approach this ideal performance.

During the training phase, we encountered an additional challenge: we only had access to the time-averaged photocurrent $\overline{I_{ph}}$ under various V_{ctr} , rather than the exact spectra of the training samples. This limitation prevents the direct application of the above method. Instead, we implemented a genetic algorithm (59) for in-situ training, inherently accounting for the physical constraints during the optimization process. Although the actual training process differs from the theoretical method outlined above, the analysis and its analogy to the kernel function remain valid, providing a solid theoretical foundation for our approach.

Error Analysis for Two Optimization Problems

We evaluated the error introduced in two optimization problems concerning the number of training samples N , the number of illuminations used J , and the number of resolved wavelengths K . This analysis evaluates the performance of the kernel function itself, focusing on how well the model separates the classes in a purely numerical sense. After that analysis, we discuss how effectively our physical system can approximate the numerically derived optimal solution.

By analyzing these errors, we can better understand the factors that influence system performance, identify the trade-offs between theoretical accuracy and physical realizability, and determine the scenarios in which our SKM is most suitable.

First Optimization Problem

The first optimization problem, based on the kernel function or support vector machine (SVM) framework, introduces some errors, which depend on various factors, such as the number of training samples N , data dimensionality (K), and class balance.

For the linear kernel discussed, a common guideline is to ensure that the number of training samples N is proportional to the number of features K , specifically, $N \geq 10K$. This rule of thumb ensures sufficient representation of the feature space for effective performance. To validate this guideline, we investigated the impact of the number of training samples on the testing score, as shown in fig. S14. We synthesized datasets with two classes of data, each containing 28-

dimensional spectral features, i.e. $\mathbf{x}_i \in \mathbb{R}^{28}$, drawn from a Gaussian distribution. Linear SVM classifiers were trained with varying training sample sizes ($N \in \{100, 150, \dots, 750, 800\}$). The testing score increased with the number of training samples, getting saturated when $N = 300 \geq 28 \times 10$. This result confirms the guideline, demonstrating that sufficient training samples are crucial for effective performance in high-dimensional feature spaces.

Ultimately, the optimal number of samples can be determined empirically by analyzing learning curves. By observing performance stabilization as the number of samples increases, we can identify the point at which additional data no longer significantly improves the model's performance.

Second Optimization Problem

When $K \leq J$, the system is undetermined, and the error is typically 0. In the more common scenario where $K > J$, Increasing K indicates approximation under more resolved wavelengths, potentially increasing the residual. To evaluate the residual errors, we initialized A , r and b in Eq. S18 according to a normal distribution $\mathcal{N}(0, 1)$ without additional constraints. The least-squares solution, denoted as \hat{r} , was computed, and the 2-norm residuals ($e = \|b - A\hat{r}\|_2$) were determined. Figure S15 displays the residuals as a percentage $\frac{e}{\|b\|_2}$ for various K and J . It can be seen that residuals increase with larger K and decrease with larger J . When $K \leq J$, the residual is 0.

We can quantitatively evaluate this behavior using the condition number of A defined as:

$$\kappa(A) = \frac{\|A\|_2}{\|A^{-1}\|_2} \quad (\text{S19})$$

A large condition number amplifies the residual:

$$\|Ar - b\|_2 \propto \kappa(A) \quad (\text{S20}).$$

For well-conditioned systems ($\kappa(A) \approx 1$), the residual remains bounded and close to zero if the system is undetermined.

Quantitative Evaluation with Hyperspectral Datasets

We further quantitatively explored the fitting error and its impact on classification accuracy with experiments (fig. S6) and simulations on hyperspectral datasets collected in the field. Figure S13E shows the example $Q(\lambda)$ highlighting class 6 (“stubble”) fitted under different spectral encoding resolutions, which is approximated with Gaussian peaks discussed in detail in Materials and Methods. The fitting error and the F_1 scores under different spectral encoding resolutions suggested a necessary spectral encoding resolution around 66 nm (32 Gaussian bands) for the tasks defined by the Salinas and Indiana Pines scenes. For higher-contrast features, such as water in the Pavias Center scene, the identification remains accurate down to 3 bands. Such results indicate that the effectiveness of the fitted kernel machine depends on factors such as the complexity of the spectral features in the task and the number of encoded spectral bands, and physically feasible spectral encoding resolutions can achieve sufficient spectral classification accuracy for real field implementations. In summary, our analyses above validate the feasibility of optoelectrically fitting the kernel machine formalism.

2. Device Metrics for Optimal SKM Performance

Having a tunable bipolar response is important to enable SKM systems. Here we provide some detailed metrics for SKM devices:

- (1) **Detectivity:** To ensure sufficient signal-to-noise ratio (SNR) of SKM at ambient, passive illumination conditions, the in-sensor computation must overcome the noise, which sets requirements on the device detectivity. Typical camera systems capture the light field and project it to the sensor array at around 0.1 W/cm^2 or below (60,61). This value corresponds to 100 nW of power illuminated on a $10 \mu\text{m} \times 10 \mu\text{m}$ device. Considering spectral encoding that reduces a significant amount of transmitted power, and an extra SNR standard on in-sensor computation, we empirically applied a 100-times lower noise equivalent power (NEP) as a requirement. The specific detectivity $D^* = \frac{\sqrt{A\Delta f}}{NEP}$ equaled to 10^9 Jones at 1 MHz operation, or $7.7 \times 10^6 \text{ Jones}$ at the video frame rate of 60 Hz. Typical silicon photoconductors and photodiodes can easily enable detectivity above 10^9 Jones in the visible band (62), well-satisfying this requirement. For MIR-related applications, room temperature MCT detectors and bP detectors can achieve $5 \sim 8 \times 10^{10} \text{ Jones}$ (37). However, these MIR devices were not tailored for low dark current and electrical tunability as

required by the SKM discussed in the manuscript. For applications inspecting objects close to the SKM, active illumination can work as the spectral encoder and provide much higher on-sensor light intensity, so that the required detectivity can be much lower, satisfying the requirement for MIR-related sensing applications.

(2) Responsivity and dark/drift current: As discussed in the main text, the dark current and its drift can significantly impact the accuracy and stability of the SKM results. Considering suppression of the dark/drift current at a level equal to the NEP photocurrent, we had $R \cdot NEP > I_{dark}$, or $\frac{R}{I_{dark}} > 10^9$ W. Photoconductors with silicon and other materials could achieve these requirements (63-65) with proper device structure engineering. An even lower dark current with photodiode structures is preferred, especially for longer wavelength photodetectors.

3. bP-MoS₂ Diode Photodetection Mechanism

Multiple factors could contribute to the bipolar photo response, including competing photo responses at bP-MoS₂ and bP-metal junctions, flipped built-in field at the vertical bP-MoS₂ interface, and flipped built-in field at the lateral bP homojunction. As discussed for fig. S9D, the photocurrent was majorly generated at the MoS₂ edge that intersects with the bP layer. Furthermore, devices with nickel, graphene, and gold contacts all show similar bipolar behavior despite different Schottky barrier heights in contact with bP, further ruling out a competing photocurrent generated at the Schottky barrier as the dominant mechanism.

The vertical heterojunction between bP and MoS₂ could also possibly dominate the bipolar photo response. Under this hypothesis, the built-in electric field between bP and MoS₂ is flipped in the vertical diode layout by the gate bias (see fig. S16). Then, the MoS₂ layer would be more intrinsic at $V_{ctr} \ll 0$ V, pushing the photoelectrons to the bP side, so $I_{ph} > 0$. However, the experimental trend was the opposite, instead supporting a homojunction mechanism (Fig. 4B). The observed trend occurred as the right half of the bP layer was screened from gating by the underlying MoS₂ layer, and the left half was subjected to the gate doping, forming a p⁺-p⁻ junction ($I_{ph} < 0$) at $V_{ctr} \ll 0$ V and an n-p⁻ junction ($I_{ph} > 0$) at $V_{ctr} \gg 0$ V. Simulated band diagrams show that the built-in electric field flips direction because the bP homojunction changes polarity under gating

(Fig. 4B). On the other hand, the bP-MoS₂ interface field does not flip direction due to strong screening by thick MoS₂.

The homojunction hypothesis was further examined by fabricating devices with different bP thicknesses. We found that bipolarity only occurred with bP thickness below ~17 nm, whereas devices with thicker bP showed a unipolar response (fig. S9C). This observation can be well explained with the homojunction model: As the exfoliated bP is naturally p-type, the thicker layers of the narrow-bandgap bP cannot be fully depleted, leading to a much weaker built-in electric field in the homojunction, allowing the bP-MoS₂ interface to dominate the photo response instead.

In conclusion, the MoS₂ layer screens half of the bP layer from the gate, creating a gate-tunable bP homojunction. MoS₂ then harvested the photo charges with low electron barrier height and high hole barrier, resulting in low dark current and good SNR for the SKM.

4. Performance Analysis and Comparisons

Here we estimated the power, latency, and data throughput of different spectral machine vision architectures, and quantitatively compared the SKM with other technologies. In Fig. 5, $J = 32, 204$, and 206 for the three datasets (with down sampling), respectively. For Fig. 6 as a generalized estimation, we assumed $J = 100$ spectral bands to be utilized for spectral machine vision, which is typical for hyperspectral imaging.

Figure 6A illustrates a conventional architecture that collects raw data for digital post-processing. Commercial products use the MIPI-CSI-2 interface to convert raw photocurrents from photodetector arrays into digital video streams. The MIPI-CSI-2 system involved analog amplification, analog-digital-conversion (ADC), and data transmission control. Sony IMX477 is a representative product using the MIPI-CSI-2 protocol, with a maximum output bandwidth of 740 MS/s at 10-bit precision. The total power consumption for the system is around 1 W, corresponding to 1.35 nJ/voxel. The latency is $100/60\text{Hz} = 1.67$ s for one collection of a full-resolution hypercube.

For digital post-processing, we assumed NVIDIA Jetson AGX Orin as the digital processor. It supports an input bandwidth of 40 GB/s, much larger than the camera output. Therefore, the sensor-processor interface is the major barrier for efficient machine vision, which is consistent with our main text discussions. Jetson AGX Orin supports 200 TOPS at 8-bit precision. Consequently, the total latency majorly comes from the input bandwidth, at 40 ms/hypercube, with

the computational power around $2 * 30 \text{ W} / 200 \text{ TOPS} = 0.3 \text{ pJ/voxel}$. Note that this estimated the upper performance limit because we assumed zero waste of the computational throughput, and no usage of CPU.

With the estimations above, we can calculate the power consumption and latency normalized to each raw voxel. For conventional architecture, power consumption majorly comes from the CSI-2 architecture that prepares the data and sends it to the digital processor, which is about **1.35 nJ/voxel**. The total latency is $\sim 1.71 \text{ s/hypercube}$ for the 4056×3040 resolutions of IMX477, with a CSI-2-limited data throughput of **0.74 Gvoxels/s**.

The SKM needed additional control, spectral encoding, and DAC modules for the in-sensor computation. We assumed a field implementation of precision agriculture. A Lyot filter can be mounted over the camera lens to parallelly encode all input spectra. The Lyot filter voltage swept in 2 ms cycles as a continuous spectral encoding over the assumed $J = 100$ bands. Since the frame rate was limited by the CSI-2 to 60 fps for full resolution, we could accumulate the photo charge for $N_{avg} = 8$ cycles of spectral encoding. Previous work indicates relatively low power consumption of the liquid-crystal-based filters. Consider a typical liquid crystal switch capacitance of $C = 0.7 \text{ nF}$ with an aperture area of 1 cm^2 , with a typical sweeping voltage range at $\Delta V = 3 \text{ V}$ (14). The Lyot filter consumes an energy of $E = Q\Delta V = C\Delta V^2 = 6.3 \text{ nJ/cycle}$. Consequently, the normalized power consumption for spectral encoding was $N_{avg}E/J/N_{pix} = 0.005 \text{ fJ/voxel}$, negligible compared to that from other modules. For wafer inspections or microscopy, LED array illuminations can be an alternative method to provide faster spectral encoding at microseconds. The light source power was below 1 W. Assuming an encoding frequency of 100 kHz and the same N_{avg} , $E = 10 \mu \text{W/cycle}$. $N_{avg}E/J/N_{pix} = 0.7 \text{ pJ/voxel}$, still significantly lower than other components.

The same conclusion applies to the encoding control and DAC modules. We assumed an Intel Core i7 CPU with $\sim 180 \text{ pJ}$ per operation. It controls a typical DAC, Texas Instruments DAC8830, which consumes $10N_{avg} \text{ pJ/Byte}$ for broadcasting the control voltage to all gated photodiodes. The power consumption on each raw voxel data reduces by 10^6 times since all spatial pixels share the same control voltage.

Lastly, we also estimated the power consumption from capacitive charging of the gated photodiodes. We assumed $10\text{ }\mu\text{m}$ pixel size and 50 nm SiO_2 as the gate dielectric. For each voxel, the weight must be separately updated. Assuming 1 V shift of control voltage between each update, as proved feasible in Fig. 4, and 8 cycles of update for averaging of one readout total photocurrent, the charging power is 0.64 pJ/voxel . Consequently, the normalized power consumption was dominated by the MIPI CSI-2 and **<15 pJ/voxel**. The total latency of the CSI-2 bottleneck was also reduced to **16.7 ms/hypercube**, with a data throughput of **74 Gvoxels/s**.

In conclusion, the SKM architecture brought about two orders of magnitude lower latency and power consumption for spectral machine vision.

5. Advantages in Practical Applications

In this section, we first discussed the speed, resolution, and power efficiency requirements for different spectral machine vision applications, which are hardly achievable with conventional spectral machine vision technologies. Next, we used numerical simulations to validate that SKM can provide sufficient accuracy for these applications.

Hyperspectral Imaging on Satellites

We first considered using earth-observing satellites for hyperspectral machine vision. The satellite requires a resolution of 1 m for resolving cars and tree canopies. Its speed relative to the ground is $\sim 10^4\text{ m/s}$, which requires tiled readout at the push-broom direction (66). With $N_{spec} = 100$ and $N_{pix} = 10,000 \times 100$ tiling covering 10 km wide push-broom, the frame rate needs to be 100 Hz for 1-m resolution. The power budget will be 13.5 W even with our lowest estimation with commercial products, or 4.5% for the average power (300 W) of earth observation satellites such as EO-1 (67), or $\sim 116\%$ of power for smaller satellites such as HYPSO-1 (49). More practically, the spectral machine vision module should consume $< 1\text{ W}$ for a small satellite system, which is equal to 0.1 nJ/voxel . The required data throughput is 10 Gvoxels/s .

Hyperspectral Machine Vision in Cellphones and Autonomous Vehicles

For snapshot hyperspectral imaging in cellphones and autonomous vehicles, consider $N_{pix} = 1280 \times 720$ (720p resolution) with $N_{spec} = 100$ at 60 fps video rate, the commercial product estimation leads to a power consumption of 6.3 W . Considering the iPhone 14 battery at 12.7 Wh ,

it is clearly challenging for long-term field applications such as precision agriculture. We set the practical power consumption to be 5%, which is the normal cellphone camera proportion (68) and considered iPhone 14 power for 8-hour usage, which is 0.08 W. The parameters correspond to 14.5 pJ/voxel and 5.5 Gvoxels/s. The power requirements for car cameras can be less strict at around 20 W, corresponding to 3.6 nJ/voxel.

Bioimaging

Spectral resolution has been proven important for resolving real-time metabolic activities in living tissues (69). These microscopic bio-tissue observations often demand high spatial resolution beyond 10 Mpixels. The required frame rate depends on the bioactivity of interest. For example, the neuron firing takes milliseconds; the bacterial flagellar rotation can be as fast as 1700 rpm, or 32 ms/cycle (70); and many biochemical processes can take minutes. Therefore, mega-pixel resolution with at least 60 fps, and 100 spectral bands are reasonable assumptions, leading to 6 Gvoxels/s data throughput. The application is less sensitive to power consumption, since the microscopy setup is typically stationary. Hence, we can assume a much larger power consumption at 10 W, corresponding to 1.7 nJ/voxel.

We plotted these performance benchmarks for satellites, cellphones, autonomous vehicles, and biomedical imagers in Figs. 6C, D, together with the performance metrics of existing commercial products. Clearly, the SKM architecture shows two orders of magnitude higher energy efficiency and data throughput, which are necessary for the above applications.

To verify the machine vision accuracy of the SKM, we modeled it in Python (see Materials and Methods), and used them to analyze different field test hyperspectral imaging datasets, including Salinas Scenes for precision agriculture, LIB-HSI for autonomous vehicles, and HS Human Brain dataset for biomedical imaging, all of which showed decent accuracy (Figs. 5D, E). For evaluating the power and latency in Figs. 5F, G, the data volume was organized to be the same for every hypercube for each dataset presented in Fig. 5.

Comparison with Geometry/Structure and RGB-Based Computer Vision Methods

In many computer vision tasks, the difference between target objects and the rest of the scene is apparent with geometrical and structural information captured with RGB sensors. But there are limitations to this approach, and in many applications, spectral domain measurements can provide

performance enhancement. There are also various tasks where pure spectral domain computation is irreplaceable by the geometry/structure or RGB-based data processing. For example, when the material features are hidden or suppressed in the invisible bands, RGB-based structural information could mislead standard machine vision approaches; see e.g., Fig. 3. Figure 5A shows another example: although the RGB rendering in Fig. 5A may give the impression that the algorithm was simply clustering pixels based on color, in many complex cases, this was not the case. The SKM operated on the full hyperspectral data cube, enabling it to identify subtle spectral differences that are not distinguishable in the RGB space. For example, the areas in the colored boxes in fig. S19 show similar RGB colors for different classes. The same conclusion held for Fig. 5C, where the tumor region could not be effectively separated using RGB channels alone, but it was clearly resolved through the spectral information captured by SKM.

Furthermore, when the spatial resolution of the image is limited, then spectral information becomes a major source of information for various machine vision-related tasks. The significance of spectral information is especially true in, for example, rescue tasks where an aircraft must inspect the landscape with a large field of view and fast sweeping speed, and therefore the size of the targets, such as plane debris, lifebuoy, or human, may only span a few pixels. Similar situations might also happen when autonomous vehicles try to identify remote objects with limited camera/ light detection and ranging (LiDAR) resolution, or street scenes with strong glares or abrupt illumination contrasts. These artifacts can degrade the performance of standard computer vision methods, and might be overcome with additional spectral information (71).

To further support these points, we conducted an additional analysis using the Salinas dataset, where we trained SKM using only the RGB channels instead of the full hyperspectral input. The RGB-based classification results present a much lower sensitivity of 66.50% compared to 89.31% for the full-spectra SKM, with an F_1 score of 60.43% compared to 73.78% originally. The results from Classes #8, #15, and #9 suggest that the lack of IR features led to significant misclassification of the visibly similar colored regions, as shown in fig. S19, which supported our conclusions discussed above.

Finally, we should emphasize that the SKM framework is compatible with spatial post-processing. The computation can be first performed in-sensor in the spectral domain, then post-processed digitally with algorithms considering geometrical factors that span multiple pixels in a scalable

SKM array. Such a joint feature extraction capability from both the spectral and spatial dimensions (see e.g., our results reported in Fig. 2K) can improve performance with reduced computational budget (71,72).

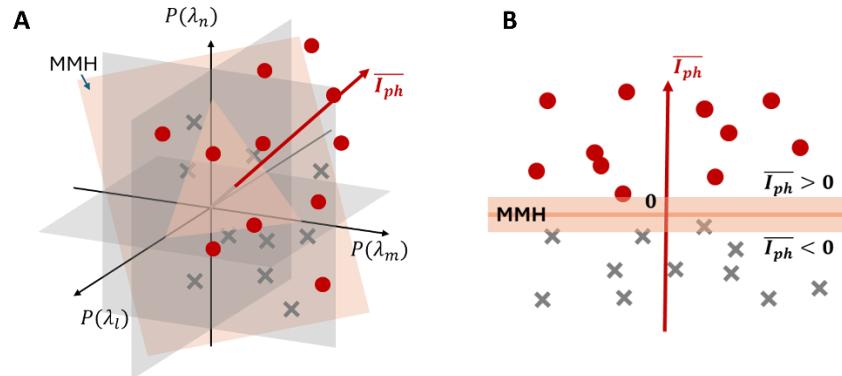


Fig. S1. SKM operation in the spectral hyperspace. (A) Red dots: positive cases (bird region in Fig. 1A). Gray crosses: the negative background. A maximum-margin hyperplane (MMH, light orange) separates the positive and negative pixels. The $\overline{I_{ph}}$ axis (red) is trained to be perpendicular to the MMH. (B) Rotated view of (A).

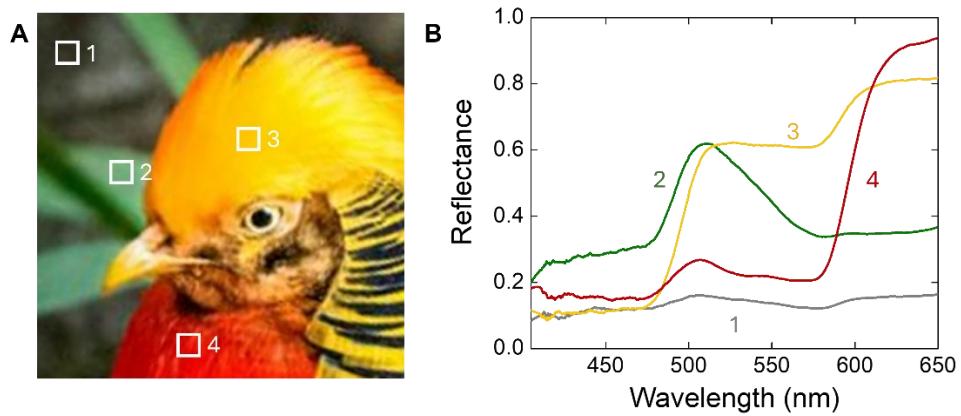


Fig. S2. Calibration spectra for the test image. (A) White boxes (with labels 1-4) mark the regions where the reflectance spectra were calibrated. (B) The measured reflectance spectra at the regions with the corresponding label numbers. The curve colors correspond to the colors of the regions.

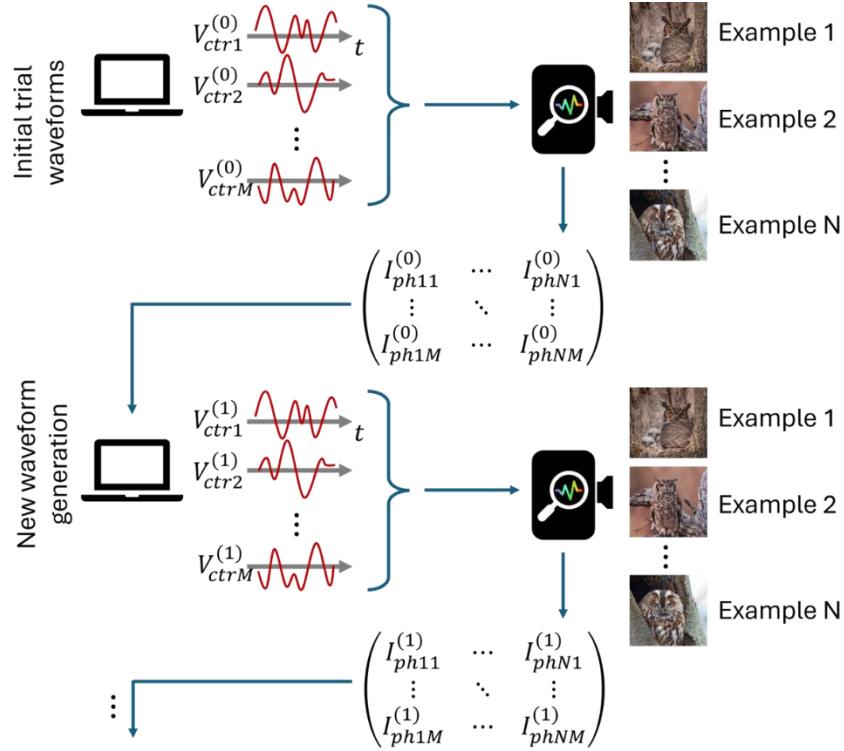


Fig. S3. Schematic of the training process for the SKM. The exampled workflow for the genetic training algorithm. The computer sends M waveforms to the sensor array, with the sensor array exposed to N example scenes, collecting an $N \times M$ photocurrent matrix for each pixel. Machine learning algorithms use these matrices to determine and generate next-round waveforms that are expected to provide more accurate spectral classification outputs from the sensor array. The performance converges after a few training cycles.

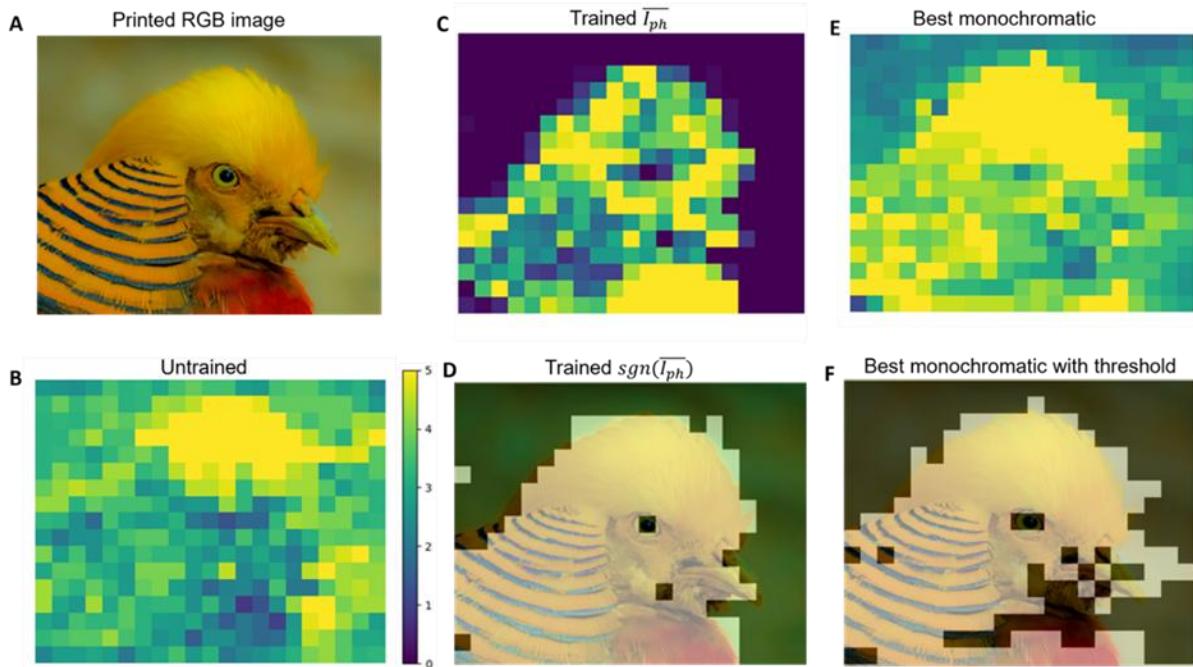


Fig. S4. Control tests for intelligent object identification. (A) The raw image. (B) The untrained SKM output. (C) The trained photocurrent. (D) The signed trained photocurrent overlapping with the original image, showing a good classification accuracy. (E) Test with the red LED as the only illumination source. The red LED is chosen since it best suppresses the background compared to the other LEDs tested. (F) The binary classification results when an optimized threshold value (3.74) is applied to E. There is a clear confusion of dim yellow and dim red regions with bright green background regions, showing that color filtering fails to perform the task properly. The comparison further demonstrates the advantages of SKM for intelligent vision.

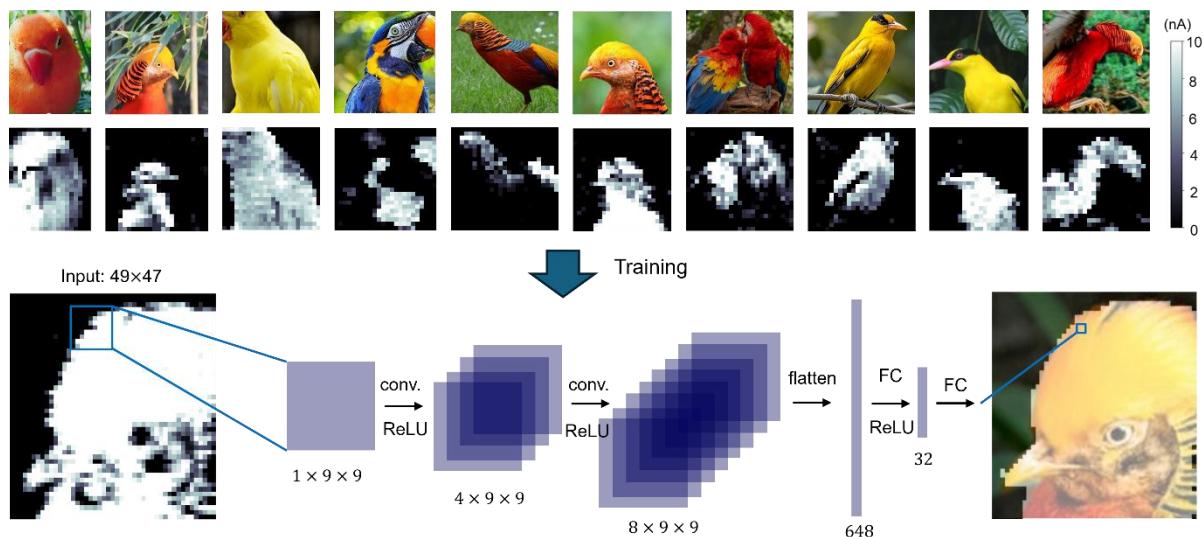


Fig. S5. SKM combined with spatial post-processing for improved image segmentation. The SKM's total photocurrent mappings from the training images (upper panel) are used to train a CNN (lower panel) to better classify the test image (never seen before). Conv: convolution layer; ReLU: rectified linear unit activation layer; FC: fully connected linear layer.

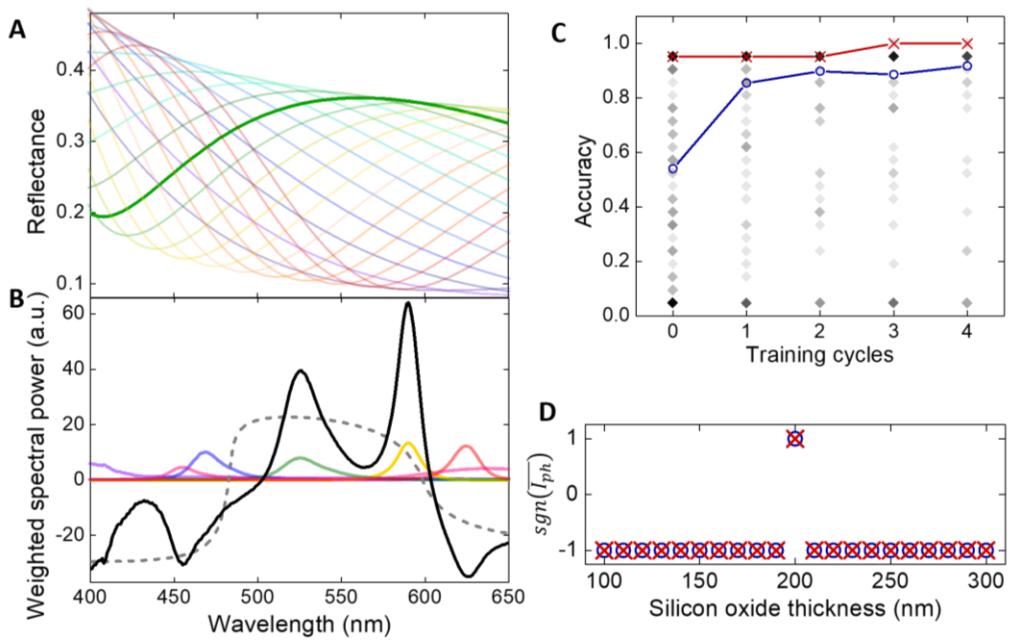


Fig. S6. SKM for binary wafer classification. (A) The reflectance of wafers at 45° incident angle. Red to violet: SiO₂ thickness at 100–300 nm with 10 nm step. Green: the positive case with 200 nm SiO₂. (B) LED spectra (colored) superposed into $Q(\lambda)$ as an approximator (black) of the ideal weighted spectrum $b(\lambda)$ (dashed gray line). (C) The learning curve with all trial solutions (diamonds), the mean (blue) and best (red) accuracy. (D) Classification results, with a good match between expectation (blue circle) and trained results (red cross).

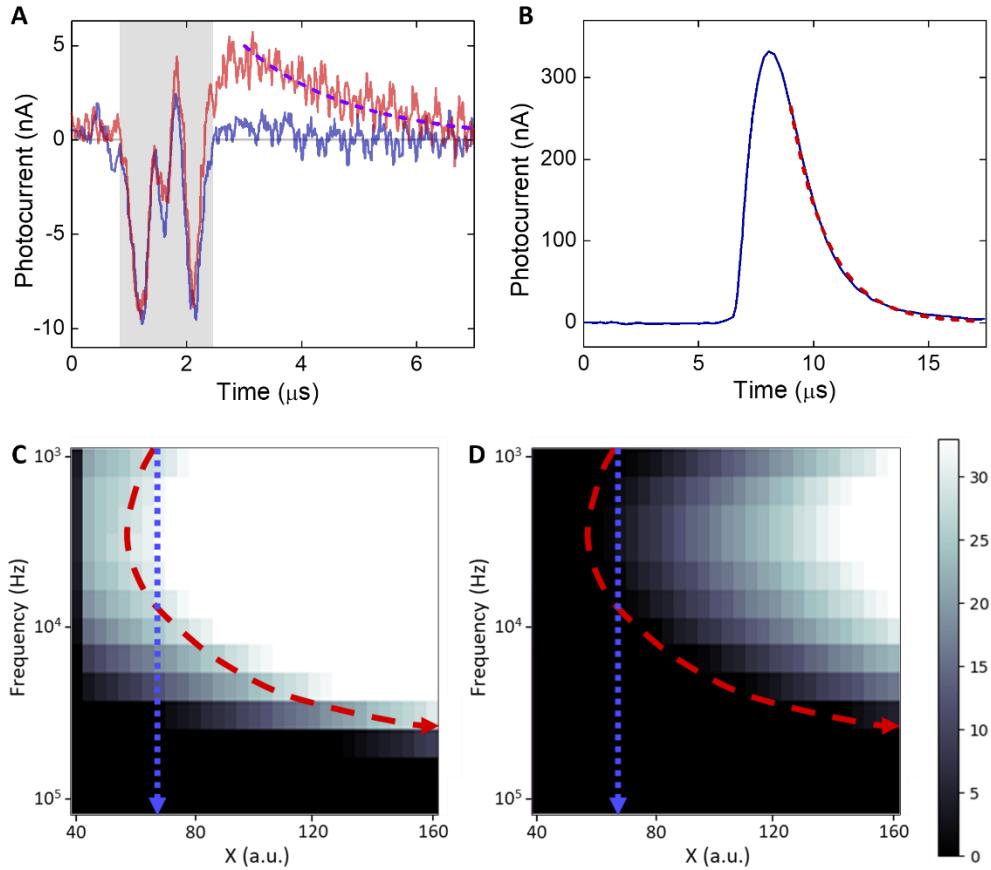


Fig. S7. Temporal response of the visible SKM. (A) Time-resolved photocurrents inspecting wafers with two oxide thicknesses: 140-nm (positive case, red) and 50-nm (negative, blue). $V_{ctr}(t)$ caused an oscillation marked by the gray zone. After that, a positive photocurrent was observed for a new positive sample for blind test, and the negative sample's photocurrent is close to zero. (B) The temporal photo response extracted from the photoconductor when an LED is modulated on/off. The system bandwidth limit at 1.65 μ s by fitting the fall slope with an exponential function. (C) The dependence of \overline{I}_{ph} (colors in grayscale) on operation frequency and V_{ctr} . X on the horizontal axis is the V_{ctr} value (between ± 0.1 V, digitized to a range of 0-255) for the blue LED illumination, other V_{ctr} values remained constant. The positive sample with 140 nm oxide was inspected. (D) The negative sample with 50 nm oxide was inspected. The blue dotted lines in (C) and (D) represent the readout with fixed weights pre-calibrated at low frequencies, which reduces to zero for both positive and negative samples at high frequencies and fails to make correct predictions. Instead, by changing X following the red dashed line, the in-situ training can compensate for the transient photo charge transport at high frequencies and still make correct predictions.

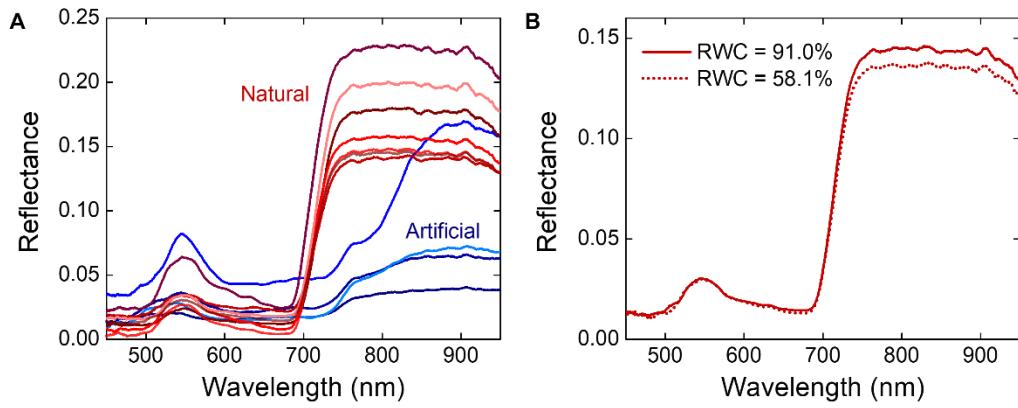


Fig. S8. Reflectance spectra for leaf-related experiments. (A) Spectra of natural (red curves) and artificial (blue curves) leaves used in the experiments reported in Figs. 3D, E of main text, showing distinctive spectral features. (B) Natural leaf spectra at different relative water concentrations (RWC), calibrated with samples corresponding to Fig. 3F. These spectra were measured with the same leaf under a high hydration level (solid line) and a low hydration level (dotted line).

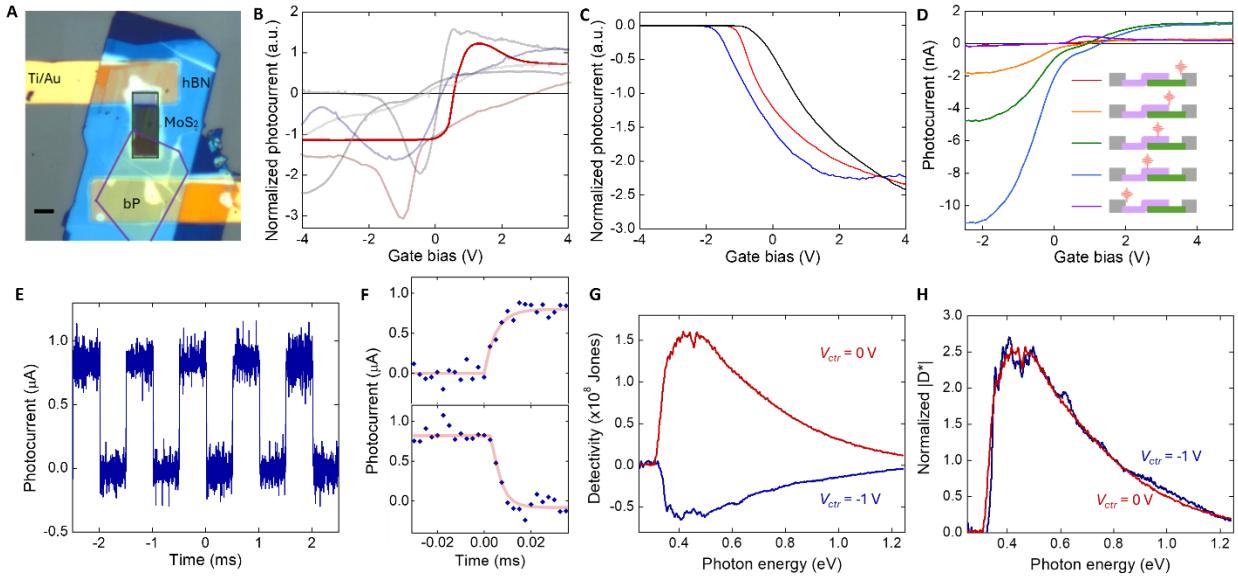


Fig. S9. Additional data for the bP-MoS₂ photodiode. (A) Optical photo of an example device. Scale bar: 5 μm . (B) The simulated bipolar response of the diode (red line). The other five curves in light shades are photo responses from different batches of devices, suggesting that the electrically tunable bipolar response is highly repeatable. (C) Unipolar photo responses from bP-MoS₂ photodiodes with thicker bP (>20 nm). The photocurrents reduce to zero at large negative gate bias because of electron concentration decrease in the MoS₂ layer, causing a larger series resistance. (D) The spatially-resolved photocurrents measured with a $\lambda = 1.62 \mu\text{m}$ laser beam focused on different spots marked in the inset. (E) Temporal response under modulated 1.55 μm laser. (F) The rise slope (upper) and fall slope (lower) fitted with exponential functions (red lines). (G) Detectivity spectrum under different gate biases, with the negative detectivity following the sign convention of the photocurrent. (H) Normalized magnitudes of the detectivities in (G).

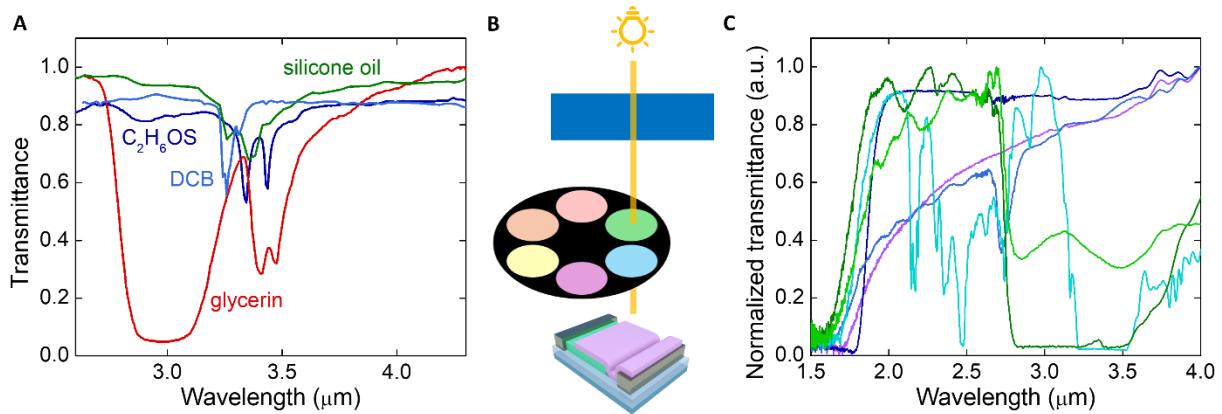


Fig. S10. NIR-MIR spectral characterization of the SKM system. (A) The transmission spectra of different test materials from the NIST online dataset. (B) The optical setup diagram. A rotating filter is synchronized with the gate voltage of the bP-MoS₂ photodiode. (C) The filter spectra calibrated via an FTIR system.

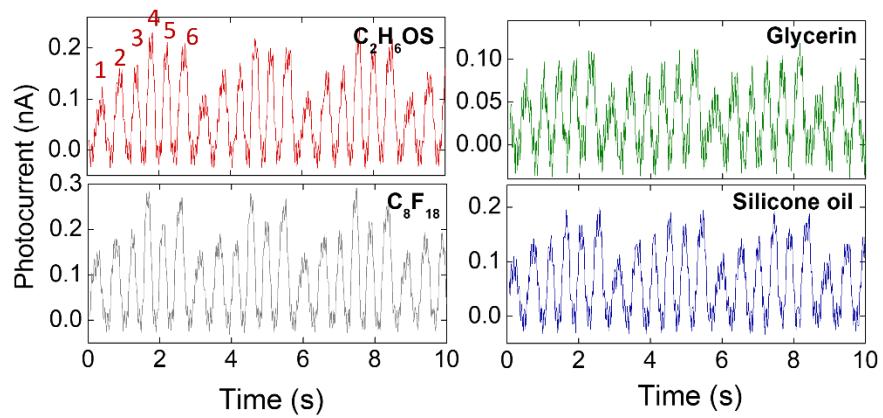


Fig. S11. Additional data for the MIR binary chemical classification. Photocurrent features in the presence of different chemicals under a rotating filter array and the black body source. The gate voltage is kept constant. The rotator is intentionally slowed down to resolve the differences visually. The peaks are aligned accordingly, showing different fingerprints for different materials, which the SKM uses to compute the classification results.

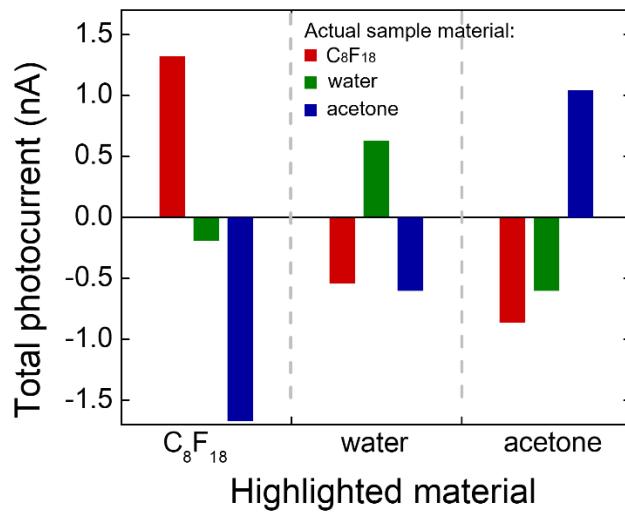


Fig. S12. Ternary classification results. The SKM is trained to classify three different materials. For example, it selectively highlights C_8F_{18} (red) and suppresses water (green) and acetone (blue) in the left part of the plot, but highlighting water/acetone in the middle/right, respectively.

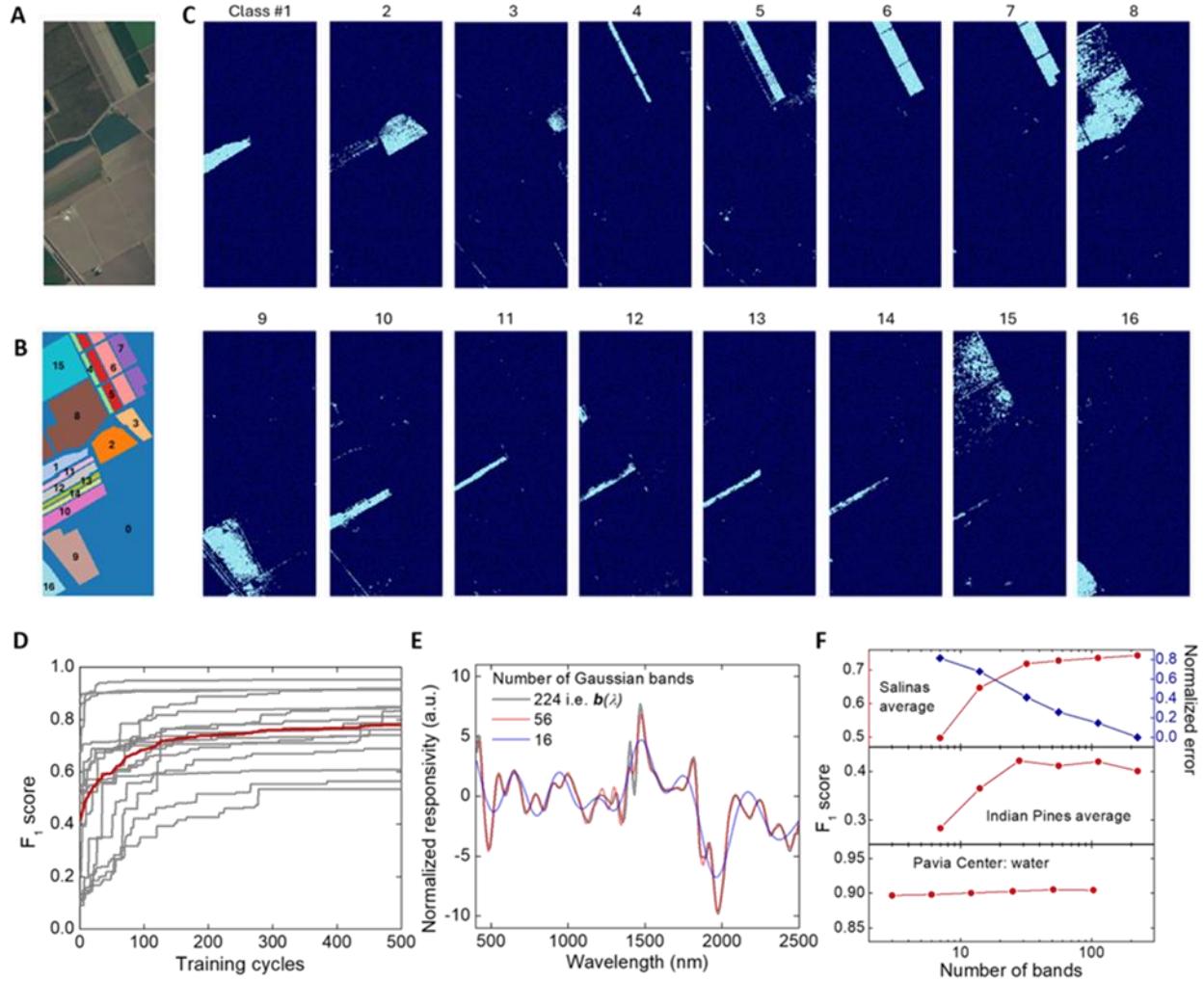


Fig. S13. Simulations on kernel fitting effectiveness. (A) The RGB image. (B) Labels of different crops/regions. (C) The binary classification results of all the 16 classes after 500 cycles of sniff. The color conventions are the same as in Fig. 5. (D) Learning curves represented by the F_1 scores of 16 classes (gray), together with the mean values (red). (E) Example $Q(\lambda)$ for highlighting class 6, fitted with different numbers of Gaussian encodings. (F) Dependence of F_1 scores and loss on the number of encoding Gaussian bands. The error (blue) is the root-mean-square (RMS) error normalized by the mean absolute value of $Q(\lambda)$ in (B). Red: the F_1 scores evaluated with three different datasets: Salinas, Indian Pines, and Pavia Center hyperspectral datasets (45). The Salinas and Indian Pines results are averaged over all classes. The Pavia Center result is for identifying water only.

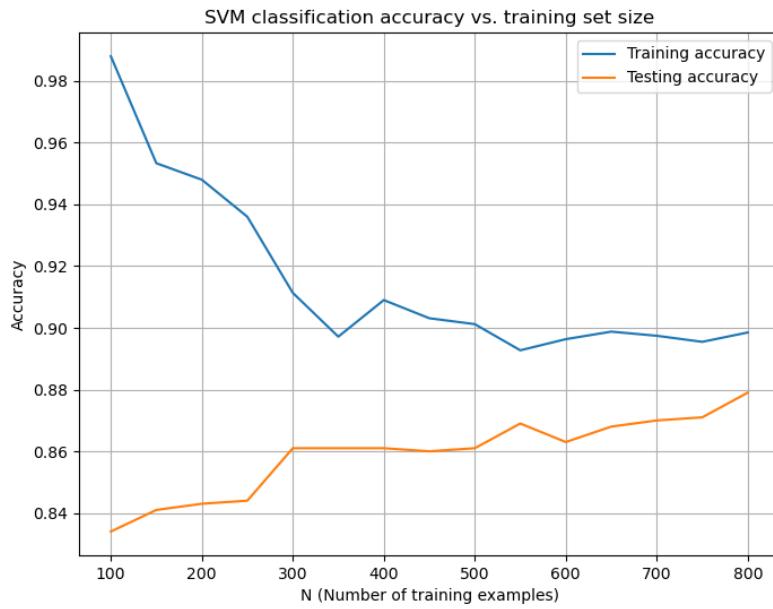


Fig. S14. Impact of the number of training samples N on the SVM classification accuracy. We used synthesized data with 28-dimensional spectral features following normal distribution, as discussed in the Supplementary Text, to investigate the impact of the number of training samples on SVM performance following Eq. (S7). 100, 150, ...800 training samples and 1000 testing samples were used for this analysis.

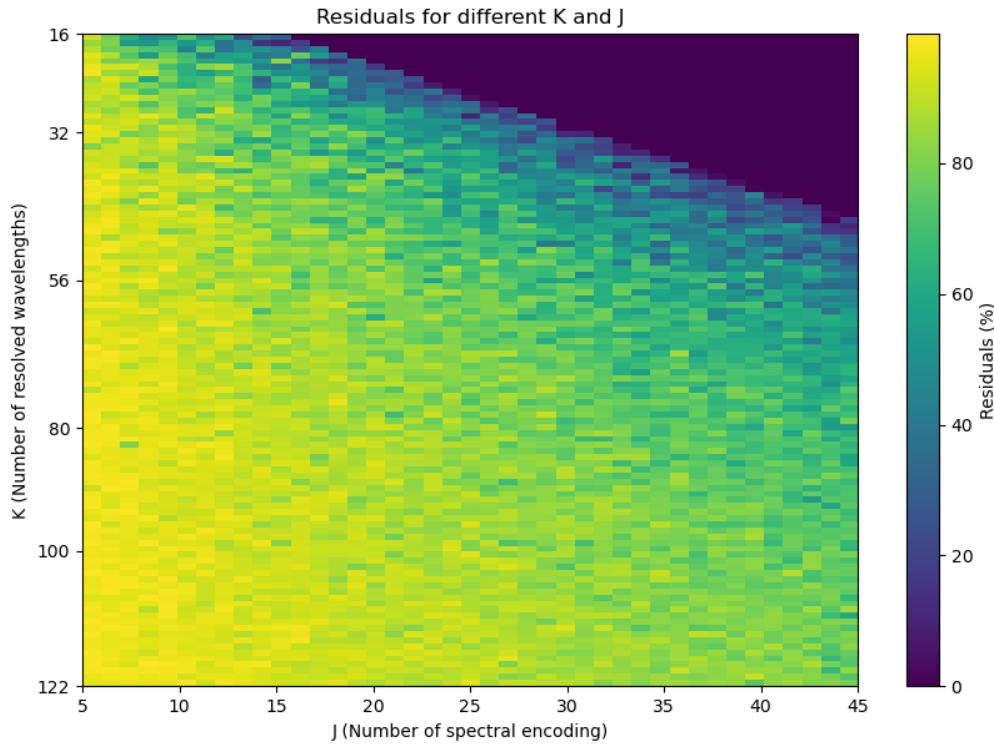


Fig. S15. Evaluation of residual errors in the second optimization problem with different dimensions K and J .
 A , r and b in Eq. S18 were initialized according to a normal distribution $\mathcal{N}(0, 1)$ without additional constraints. The least-squares solution, denoted as \hat{r} , was computed, and the 2-norm residuals ($e = \|b - A\hat{r}\|_2$) were determined. The residuals as a percentage $\frac{e}{\|b\|_2}$ for various K and J were displayed.

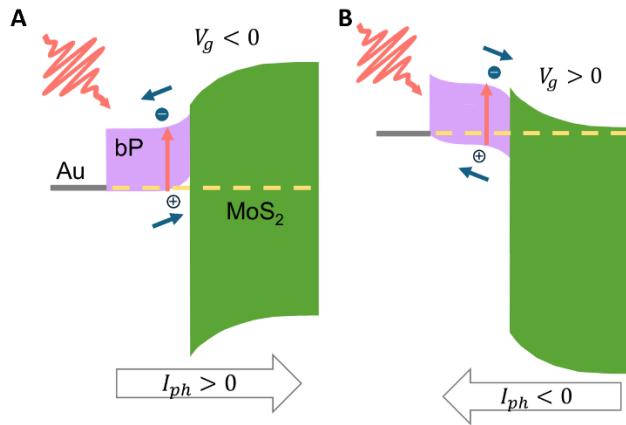


Fig. S16. Band-diagram analysis of competing photodetection mechanisms. (A) Considerations on the case that the photo response is dominated by the vertical built-in electric field between the bP and MoS₂ layers. Under negative gate voltage ($V_g \ll 0$), the gate doping depletes MoS₂ and could lead to a built-in electric field pointing to the left, such model suggests a positive photocurrent. (B) The MoS₂ layer is naturally n-doped with zero or positive gating, leading to a negative photocurrent in the vertical band-dominating model.

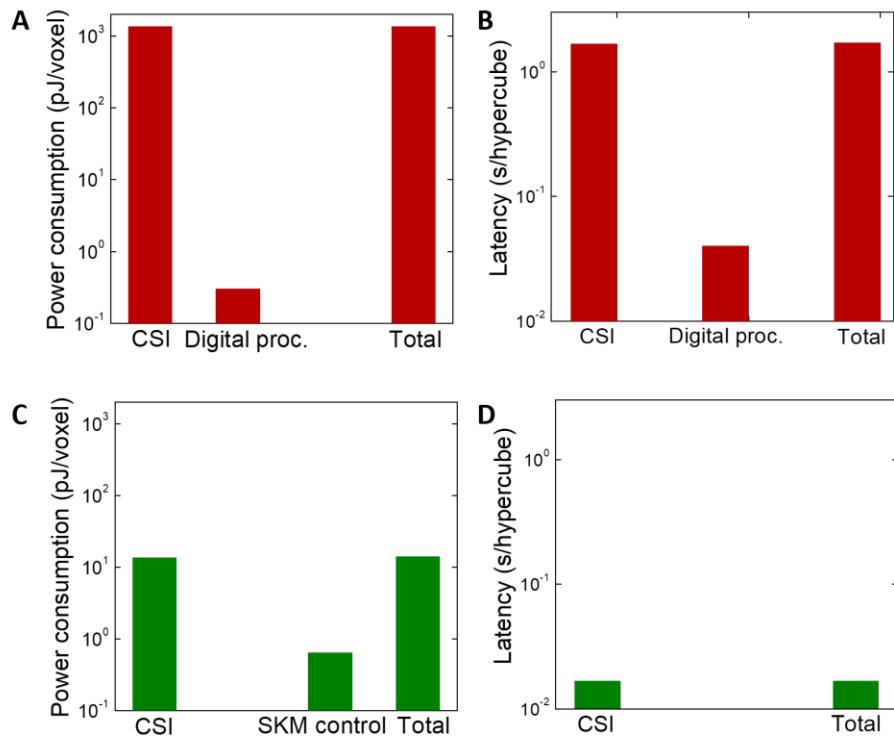


Fig. S17. Power consumption and latency contributed by different modules. (A) Power consumption of conventional digital post-processing illustrated in Fig. 6A. (B) The corresponding latency. (C) Power consumption of SKM architecture illustrated in Fig. 6B. (D) The corresponding latency. Here we assume the hypercubes have $J = 100$ spectral bands, and are with full spatial resolution of the hardware.

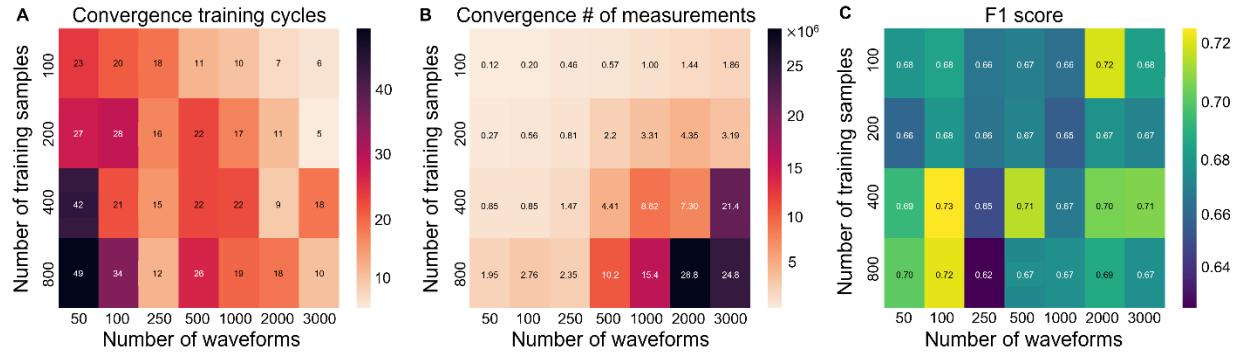


Fig. S18. Influence of the numbers of waveforms and training samples on the SKM performance. (A) The cycles required for reaching convergence. (B) The total number of measurements on various pixel-waveform combinations for reaching convergence. (C) The F_1 scores for different numbers of waveforms and training samples used to train the SKM device. The analysis was performed on the Salinas hyperspectral dataset, and the results shown are averaged across all data classes.

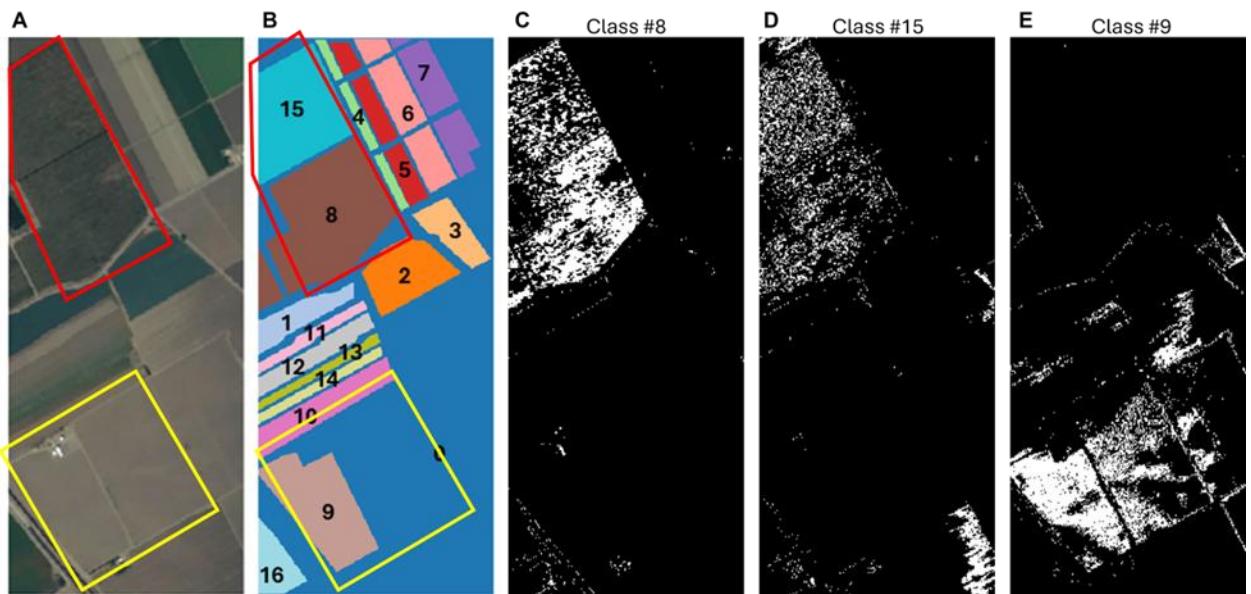


Fig. S19. SKM classification with RGB information only. (A) Two example pairs of regions (red and yellow boxes) where the RGB colors are similar but belong to different classes. (B) The ground truth. (C) The classification results for Class #8 using RGB information only. (D) Results for Class #15, which mixes with Class #8 based on RGB information only. (E) Class #9 mixing with the background region Class #0 using RGB information only.

Ref.	Type of intelligence	Auto-learning?	Spectral resolution	Spectral range	Additional notes
(10-12)	Spatial convolution	N	N	Visible	
(13-17)	Intensity adaptation	Y	N	Visible	
(18)	Color-intensity adaptation	N	RGB and white	Visible	Software postprocessing
(19)	Adaptive detection band selection	N	Two bands	Vis-IR $\sim 1.2 \mu\text{m}$	Software postprocessing for imaging application
(20)	Spectral classification	N	RGB	Visible	Sacrifice spatial resolution for spectral resolution.
(21)	Spectral classification	N	3 bands	Visible	
This work	Spectral classification, inference, and image segmentation at high accuracy	Y	6 bands (experimental) 32 bands (simulation)	Visible to MIR ($\sim 4 \mu\text{m}$)	

Table S1. Summary of different reports of in-sensor computing with neuromorphic photodetectors. The references are numbered according to the main text reference list.

	Task	Training objects	Test objects	# genes/bands	# trial $V_{ctr}(t)$	# parents mating	# parents kept	Mut. rate	Fitness function
Silicon-band SKM	Binary classification	240 random pixels from 10 images of birds in forests	A new image of bird in the forest	6	50	10	10	30%	$\frac{e^{\beta I_+}}{e^{\beta I_+} + e^{\beta I_-}} - \gamma \max(I_-, 0)$ (see Materials and Methods)
		Two silicon wafers with 90.5 nm oxide (16 points) and 90.0 nm oxide (16 points)	32 testing points (50% positive) on two new wafers	6	50	10	10	30%	
		18 natural leaves from 7 species, 18 artificial leaves from 4 manufacturers	31 natural and artificial leaves	6	50	10	10	30%	
		12 high-hydration-level leaves and 12 low-hydration-level leaves	24 new leaves (50% positive)	6	50	10	10	30%	
	Inference	Random regions on 4 wafers with 90, 105, 120, and 140 nm oxide (1 sampling/wafer/sniff)	Random regions on 10 wafers with different oxide thicknesses	6	100	10	10	30%	Negative of covariance: $-\text{cov}(\bar{I}_{ph} - t)$, t is the ground truth thickness in nm. \bar{I}_{ph} in nA
MIR SKM	Binary classification	Two chemicals in Petri dish, randomly placed	Two chemicals in Petri dish, randomly placed	6	40	10	10	30%	$\frac{e^{\beta I_+}}{e^{\beta I_+} + e^{\beta I_-}} - \gamma \max(I_-, 0)$
	Inference	Five concentrations: 0%, 10%, 20%, 30%, and 40%, beaker randomly placed	16 concentrations, beaker randomly placed	6	50	10	10	30%	Negative of covariance: $-\text{cov}(\bar{I}_{ph} - c)$, c is the ground truth concentration in percent. \bar{I}_{ph} in pA
Simulated HSI	Binary classification	10% of the whole dataset, randomly sampled	90% of the whole dataset	32	3000	1000	300	70%	F1 score

Table S2. Summary of training and test conditions. See the Materials and Methods for additional details.

Ref.	Year	Spatial Resolution	Number of bands	Frame Rate (fps)	Total power	Power consumption	Data throughput (voxels/s)	Other notes
(2)	2023	640 × 480	20	206.6	> 4.2 W	2.9 nJ/voxel	1.43 Gvoxels/s	Number of bands from spectral reconstruction resolution Assuming GPU power consumption negligible after normalization to voxel
		1280 × 720	20	76.9				
		1920 × 1080	20	34.4				
(47)	2019	1024	224	200	~12 W	260 nJ/voxel	46 Mvoxels/s	Pushbroom, power estimated from specs of spectrometer and GPU applied.
(48)	2020	1024	224	330	17 W	225 nJ/voxel	76 Mvoxels/s	
(49)	2021	1936×1194 voxels		47	3.5 W	32 nJ/voxel	109 Mvoxels/s	
(50)	2022	825×1024	160	2	1.3 W	4.8 nJ/voxel	270 Mvoxels/s	
(51)	2025	288 × 288	43	12	6.3 W	147 nJ/voxel	43 Mvoxel/s	
(52)	2024	1505 × 1764		28	>4.8 W	>65 nJ/voxel	74 Mvoxels/s	Additional post-processing power not included
(53)	2024	128 × 128	64	4	12 J /hypercube	2.9 uJ/voxel	4.2 Mvoxels/s	Power normalized by GPU usage for decoding each hypercube
Estimated best post-processing benchmark	2025	4056 × 3040	60		≥1 W	≥1.35 nJ/voxel	740 Mvoxels/s	
SKM architecture	2025	4056 × 3040	100	60	1 W	14 pJ/voxel	74 Gvoxels/s	

Table S3. Performance summary of different implementations of hyperspectral imaging and analysis.

Movie S1. Classification results during 500 cycles of SKM training performed on the Salinas hyperspectral dataset.

References and Notes

1. L. Bian, Z. Wang, Y. Zhang, L. Li, Y. Zhang, C. Yang, W. Fang, J. Zhao, C. Zhu, Q. Meng, X. Peng, J. Zhang, A broadband hyperspectral image sensor with high spatio-temporal resolution. *Nature* **635**, 73–81 (2024). [doi:10.1038/s41586-024-08109-1](https://doi.org/10.1038/s41586-024-08109-1) [Medline](#)
2. M. Yako, Y. Yamaoka, T. Kiyohara, C. Hosokawa, A. Noda, K. Tack, N. Spooren, T. Hirasawa, A. Ishikawa, Video-rate hyperspectral camera based on a CMOS-compatible random array of Fabry–Pérot filters. *Nat. Photonics* **17**, 218–223 (2023). [doi:10.1038/s41566-022-01141-5](https://doi.org/10.1038/s41566-022-01141-5)
3. A. McClung, S. Samudrala, M. Torfeh, M. Mansouree, A. Arbabi, Snapshot spectral imaging with parallel metasystems. *Sci. Adv.* **6**, eabc7646 (2020). [doi:10.1126/sciadv.abc7646](https://doi.org/10.1126/sciadv.abc7646) [Medline](#)
4. J. Xiong, X. Cai, K. Cui, Y. Huang, J. Yang, H. Zhu, W. Li, B. Hong, S. Rao, Z. Zheng, S. Xu, Y. He, F. Liu, X. Feng, W. Zhang, Dynamic brain spectrum acquired by a real-time ultraspectral imaging chip with reconfigurable metasurfaces. *Optica* **9**, 461–468 (2022). [doi:10.1364/OPTICA.440013](https://doi.org/10.1364/OPTICA.440013)
5. W. Zhang, J. Suo, K. Dong, L. Li, X. Yuan, C. Pei, Q. Dai, Handheld snapshot multi-spectral camera at tens-of-megapixel resolution. *Nat. Commun.* **14**, 5043 (2023). [doi:10.1038/s41467-023-40739-3](https://doi.org/10.1038/s41467-023-40739-3) [Medline](#)
6. R. Gebbers, V. I. Adamchuk, Precision agriculture and food security. *Science* **327**, 828–831 (2010). [doi:10.1126/science.1183899](https://doi.org/10.1126/science.1183899) [Medline](#)
7. M. T. Eismann, A. D. Stocker, N. M. Nasrabadi, Automated hyperspectral cueing for civilian search and rescue. *Proc. IEEE* **97**, 1031–1055 (2009). [doi:10.1109/JPROC.2009.2013561](https://doi.org/10.1109/JPROC.2009.2013561)
8. Y. Jiang, S. Karpf, B. Jalali, Time-stretch LiDAR as a spectrally scanned time-of-flight ranging camera. *Nat. Photonics* **14**, 14–18 (2020). [doi:10.1038/s41566-019-0548-6](https://doi.org/10.1038/s41566-019-0548-6)
9. F. M. Padilla, R. de Souza, M. T. Pena-Fleitas, R. Grasso, M. Gallardo, R. B. Thompson, Influence of time of day on measurement with chlorophyll meters and canopy reflectance sensors of different crop N status. *Precis. Agric.* **20**, 1087–1106 (2019). [doi:10.1007/s11119-019-09641-1](https://doi.org/10.1007/s11119-019-09641-1)
10. L. Mennel, J. Symonowicz, S. Wachter, D. K. Polyushkin, A. J. Molina-Mendoza, T. Mueller, Ultrafast machine vision with 2D material neural network image sensors. *Nature* **579**, 62–66 (2020). [doi:10.1038/s41586-020-2038-x](https://doi.org/10.1038/s41586-020-2038-x) [Medline](#)
11. C. Y. Wang, S. J. Liang, S. Wang, P. Wang, Z. Li, Z. Wang, A. Gao, C. Pan, C. Liu, J. Liu, H. Yang, X. Liu, W. Song, C. Wang, B. Cheng, X. Wang, K. Chen, Z. Wang, K. Watanabe, T. Taniguchi, J. J. Yang, F. Miao, Gate-tunable van der Waals heterostructure for reconfigurable neural network vision sensor. *Sci. Adv.* **6**, eaba6173 (2020). [doi:10.1126/sciadv.aba6173](https://doi.org/10.1126/sciadv.aba6173) [Medline](#)
12. L. Pi, P. Wang, S. J. Liang, P. Luo, H. Wang, D. Li, Z. Li, P. Chen, X. Zhou, F. Miao, T. Zhai, Broadband convolutional processing using band-alignment-tunable heterostructures. *Nat. Electron.* **5**, 248–254 (2022). [doi:10.1038/s41928-022-00747-5](https://doi.org/10.1038/s41928-022-00747-5)

13. F. Liao, Z. Zhou, B. J. Kim, J. Chen, J. Wang, T. Wan, Y. Zhou, A. T. Hoang, C. Wang, J. Kang, J. H. Ahn, Y. Chai, Bioinspired in-sensor visual adaptation for accurate perception. *Nat. Electron.* **5**, 84–91 (2022). [doi:10.1038/s41928-022-00713-1](https://doi.org/10.1038/s41928-022-00713-1)
14. D. Zhang, D. Xu, Y. Li, Y. Luo, J. Hu, J. Zhou, Y. Zhang, B. Zhou, P. Wang, X. Li, B. Bai, H. Ren, L. Wang, A. Zhang, M. Jarrahi, Y. Huang, A. Ozcan, X. Duan, Broadband nonlinear modulation of incoherent light using a transparent optoelectronic neuron array. *Nat. Commun.* **15**, 2433 (2024). [doi:10.1038/s41467-024-46387-5](https://doi.org/10.1038/s41467-024-46387-5) [Medline](#)
15. Z. He, H. Shen, D. Ye, L. Xiang, W. Zhao, J. Ding, F. Zhang, C. Di, D. Zhu, An organic transistor with light intensity-dependent active photoadaptation. *Nat. Electron.* **4**, 522–529 (2021). [doi:10.1038/s41928-021-00615-8](https://doi.org/10.1038/s41928-021-00615-8)
16. W. Deng, X. Zhang, R. Jia, L. Huang, X. Zhang, J. Jie, Organic molecular crystal-based photosynaptic devices for an artificial visual-perception system. *NPG Asia Mater.* **11**, 77 (2019). [doi:10.1038/s41427-019-0182-2](https://doi.org/10.1038/s41427-019-0182-2)
17. L. Li, S. Li, W. Wang, J. Zhang, Y. Sun, Q. Deng, T. Zheng, J. Lu, W. Gao, M. Yang, H. Wang, Y. Pan, X. Liu, Y. Yang, J. Li, N. Huo, Adaptative machine vision with microsecond-level accurate perception beyond human retina. *Nat. Commun.* **15**, 6261 (2024). [doi:10.1038/s41467-024-50488-6](https://doi.org/10.1038/s41467-024-50488-6) [Medline](#)
18. Y. Liu, Z. Ji, G. Cen, H. Sun, H. Wang, C. Zhao, Z. L. Wang, W. Mai, Perovskite-based color camera inspired by human visual cells. *Light Sci. Appl.* **12**, 43 (2023). [doi:10.1038/s41377-023-01072-y](https://doi.org/10.1038/s41377-023-01072-y) [Medline](#)
19. B. Ouyang, J. Wang, G. Zeng, J. Yan, Y. Zhou, X. Jiang, B. Shao, Y. Chai, Bioinspired in-sensor spectral adaptation for perceiving spectrally distinctive features. *Nat. Electron.* **7**, 705–713 (2024). [doi:10.1038/s41928-024-01208-x](https://doi.org/10.1038/s41928-024-01208-x)
20. Z. Peng, L. Tong, W. Shi, L. Xu, X. Huang, Z. Li, X. Yu, X. Meng, X. He, S. Lv, G. Yang, H. Hao, T. Jiang, X. Miao, L. Ye, Multifunctional human visual pathway-replicated hardware based on 2D materials. *Nat. Commun.* **15**, 8650 (2024). [doi:10.1038/s41467-024-52982-3](https://doi.org/10.1038/s41467-024-52982-3) [Medline](#)
21. D. Kwak, D. K. Polyushkin, T. Mueller, In-sensor computing using a MoS₂ photodetector with programmable spectral responsivity. *Nat. Commun.* **14**, 4264 (2023). [doi:10.1038/s41467-023-40055-w](https://doi.org/10.1038/s41467-023-40055-w) [Medline](#)
22. J. Ma, Q. Gu, Y. Liu, J. Lai, P. Yu, X. Zhuo, Z. Liu, J. H. Chen, J. Feng, D. Sun, Nonlinear photoresponse of type-II Weyl semimetals. *Nat. Mater.* **18**, 476–481 (2019). [doi:10.1038/s41563-019-0296-5](https://doi.org/10.1038/s41563-019-0296-5) [Medline](#)
23. W. Liu, J. Lv, L. Peng, H. Guo, C. Liu, Y. Liu, W. Li, L. Li, L. Liu, P. Wang, S. C. Bodepudi, K. Shehzad, G. Hu, K. Liu, Z. Sun, T. Hasan, Y. Xu, X. Wang, C. Gao, B. Yu, X. Duan, Graphene charge-injection photodetectors. *Nat. Electron.* **5**, 281–288 (2022). [doi:10.1038/s41928-022-00755-5](https://doi.org/10.1038/s41928-022-00755-5)
24. D. Zhang, Z. Xu, G. Cheng, Z. Liu, A. R. Gutierrez, W. Zang, T. B. Norris, Z. Zhong, Strongly enhanced THz generation enabled by a graphene hot-carrier fast lane. *Nat. Commun.* **13**, 6404 (2022). [doi:10.1038/s41467-022-34170-3](https://doi.org/10.1038/s41467-022-34170-3) [Medline](#)

25. A. Aizerman, Theoretical foundations of the potential function method in pattern recognition learning. *Autom. Remote Control* **25**, 821–837 (1964).
26. T. Hofmann, B. Schölkopf, A. J. Smola, Kernel methods in machine learning. *Ann. Stat.* **36**, 1171–1220 (2008). [doi:10.1214/009053607000000677](https://doi.org/10.1214/009053607000000677)
27. S. Pai, Z. Sun, T. W. Hughes, T. Park, B. Bartlett, I. A. D. Williamson, M. Minkov, M. Milanizadeh, N. Abebe, F. Morichetti, A. Melloni, S. Fan, O. Solgaard, D. A. B. Miller, Experimentally realized in situ backpropagation for deep learning in photonic neural networks. *Science* **380**, 398–404 (2023). [doi:10.1126/science.adc8450](https://doi.org/10.1126/science.adc8450) [Medline](#)
28. A. Momeni, B. Rahmani, B. Scellier, L. G. Wright, P. L. McMahon, C. C. Wanjura, Y. Li, A. Skalli, N. G. Berloff, T. Onodera, I. Oguz, F. Morichetti, P. del Hougne, M. Le Gallo, A. Sebastian, A. Mirhoseini, C. Zhang, D. Markovic, D. Brunner, C. Moser, S. Gigan, F. Marquardt, A. Ozcan, J. Grollier, A. J. Liu, D. Psaltis, A. Alù, R. Fleury, Training of physical neural networks. *Nature* **645**, 53–561 (2025). [doi:10.1038/s41586-025-09384-2](https://doi.org/10.1038/s41586-025-09384-2) [Medline](#)
29. A. F. Gad, Pygad: An intuitive genetic algorithm python library. *Multimedia Tools Appl.* **83**, 58029–58042 (2024). [doi:10.1007/s11042-023-17167-y](https://doi.org/10.1007/s11042-023-17167-y)
30. D. Rodrigo, O. Limaj, D. Janner, D. Etezadi, F. Javier García de Abajo, V. Pruneri, H. Altug, Mid-infrared plasmonic biosensing with graphene. *Science* **349**, 165–168 (2015). [doi:10.1126/science.aab2051](https://doi.org/10.1126/science.aab2051) [Medline](#)
31. S. M. Park, Y. Ji, S. Kwon, J. W. Leem, A. R. W. O'Brien, Y. Wang, Y. L. Kim, Remote Blood Hemoglobin Monitoring with Hyperspectral Color Truthing for Advancing Sickle Cell Care. *Blood* **142** (suppl. 1), 2277 (2023). [doi:10.1182/blood-2023-190659](https://doi.org/10.1182/blood-2023-190659)
32. D. Jia, Y. Zhang, Q. Yang, Y. Xue, Y. Tan, Z. Guo, M. Zhang, L. Tian, J.-X. Cheng, J. X. Cheng, 3D chemical imaging by fluorescence-detected mid-infrared photothermal Fourier light field microscopy. *Chem. Biomed. Imaging* **1**, 260–267 (2023). [doi:10.1021/cbmi.3c00022](https://doi.org/10.1021/cbmi.3c00022) [Medline](#)
33. J. Wang, X. Chen, W. Hu, L. Wang, W. Lu, F. Xu, J. Zhao, Y. Shi, R. Ji, Amorphous HgCdTe infrared photoconductive detector with high detectivity above 200 K. *Appl. Phys. Lett.* **99**, 113508 (2011). [doi:10.1063/1.3638459](https://doi.org/10.1063/1.3638459)
34. C. H. Liu, Y. C. Chang, T. B. Norris, Z. Zhong, Graphene photodetectors with ultra-broadband and high responsivity at room temperature. *Nat. Nanotechnol.* **9**, 273–278 (2014). [doi:10.1038/nnano.2014.31](https://doi.org/10.1038/nnano.2014.31) [Medline](#)
35. D. Zhang, Z. Xu, Z. Huang, A. R. Gutierrez, C. J. Blocker, C. H. Liu, M. B. Lien, G. Cheng, Z. Liu, I. Y. Chun, J. A. Fessler, Z. Zhong, T. B. Norris, Neural network based 3D tracking with a graphene transparent focal stack imaging system. *Nat. Commun.* **12**, 2413 (2021). [doi:10.1038/s41467-021-22696-x](https://doi.org/10.1038/s41467-021-22696-x) [Medline](#)
36. Q. Guo, A. Pospischil, M. Bhuiyan, H. Jiang, H. Tian, D. Farmer, B. Deng, C. Li, S. J. Han, H. Wang, Q. Xia, T. P. Ma, T. Mueller, F. Xia, Black phosphorus mid-infrared photodetectors with high gain. *Nano Lett.* **16**, 4648–4655 (2016). [doi:10.1021/acs.nanolett.6b01977](https://doi.org/10.1021/acs.nanolett.6b01977) [Medline](#)

37. M. Amani, E. Regan, J. Bullock, G. H. Ahn, A. Javey, Mid-wave infrared photoconductors based on black phosphorus-arsenic alloys. *ACS Nano* **11**, 11724–11731 (2017). [doi:10.1021/acsnano.7b07028](https://doi.org/10.1021/acsnano.7b07028) [Medline](#)
38. C. R. Pidgeon, C. M. Ciesla, B. N. Murdin, Suppression of non-radiative processes in semiconductor mid-infrared emitters and detectors. *Prog. Quantum Electron.* **21**, 361–419 (1997). [doi:10.1016/S0079-6727\(97\)00012-8](https://doi.org/10.1016/S0079-6727(97)00012-8)
39. S. Wang, N. Higashitarumizu, B. Sari, M. C. Scott, A. Javey, Quantitative Mid-infrared Photoluminescence Characterization of Black Phosphorus-Arsenic Alloys. *ACS Nano* **18**, 5907–5914 (2024). [doi:10.1021/acsnano.3c12927](https://doi.org/10.1021/acsnano.3c12927) [Medline](#)
40. N. Higashitarumizu, S. Z. Uddin, D. Weinberg, N. S. Azar, I. K. M. Reaz Rahman, V. Wang, K. B. Crozier, E. Rabani, A. Javey, Anomalous thickness dependence of photoluminescence quantum yield in black phosphorous. *Nat. Nanotechnol.* **18**, 507–513 (2023). [doi:10.1038/s41565-023-01335-0](https://doi.org/10.1038/s41565-023-01335-0) [Medline](#)
41. J. Bullock, M. Amani, J. Cho, Y. Z. Chen, G. H. Ahn, V. Adinolfi, V. R. Shrestha, Y. Gao, K. B. Crozier, Y. L. Chueh, A. Javey, Polarization-resolved black phosphorus/molybdenum disulfide mid-wave infrared photodiodes with high detectivity at room temperature. *Nat. Photonics* **12**, 601–607 (2018). [doi:10.1038/s41566-018-0239-8](https://doi.org/10.1038/s41566-018-0239-8)
42. https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes [Accessed online on Oct 12, 2024]
43. N. Habili, E. Kwan, W. Li, C. Webers, J. Oorloff, M. A. Armin, L. Petersson, A hyperspectral and RGB dataset for building façade segmentation. *European Conference on Computer Vision* 258–267, (2022).
44. H. Fabelo, S. Ortega, A. Szolna, D. Bulters, J. F. Piñeiro, S. Kabwama, A. J. O’Shanahan, H. Bulstrode, S. Bisshopp, B. R. Kiran, D. Ravi, R. Lazcano, D. Madroñal, C. Sosa, C. Espino, M. Marquez, M. D. L. L. Plaza, R. Camacho, D. Carrera, M. Hernández, G. M. Callicó, J. M. Molina, B. Stanciulescu, G. Z. Yang, R. Salvador, E. Juárez, C. Sanz, R. Sarmiento, In-vivo hyperspectral human brain image database for brain cancer detection. *IEEE Access* **7**, 39098–39116 (2019). [doi:10.1109/ACCESS.2019.2904788](https://doi.org/10.1109/ACCESS.2019.2904788)
45. F. Melgani, L. Bruzzone, Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. Geosci. Remote Sens.* **42**, 1778–1790 (2004). [doi:10.1109/TGRS.2004.831865](https://doi.org/10.1109/TGRS.2004.831865)
46. D. Hong, Z. Han, J. Yao, L. Gao, B. Zhang, A. Plaza, SpectralFormer: Rethinking hyperspectral image classification with transformers. *IEEE Trans. Geosci. Remote Sens.* **60**, 5518615 (2021). [doi:10.1109/TGRS.2021.3130716](https://doi.org/10.1109/TGRS.2021.3130716)
47. P. Horstrand, R. Guerra, A. Rodríguez, M. Díaz, S. López, J. F. López, A UAV platform based on a hyperspectral sensor for image capturing and on-board processing. *IEEE Access* **7**, 66919–66938 (2019). [doi:10.1109/ACCESS.2019.2913957](https://doi.org/10.1109/ACCESS.2019.2913957)
48. SPECIM AFX10, <https://www.specim.com/products/specim-afx10/>
49. M. E. Grøtte, R. Birkeland, E. Honoré-Livermore, S. Bakken, J. L. Garrett, E. F. Prentice, F. Sigernes, M. Orlandic, J. T. Gravdahl, T. A. Johansen, Ocean color hyperspectral remote

- sensing with high resolution and low latency—The HYPSO-1 CubeSat mission. *IEEE Trans. Geosci. Remote Sens.* **60**, 1–19 (2021). [doi:10.1109/TGRS.2021.3080175](https://doi.org/10.1109/TGRS.2021.3080175)
50. J. Caba, M. Díaz, J. Barba, R. Guerra, S. Escolar, S. López, Low-power hyperspectral anomaly detector implementation in cost-optimized FPGA devices. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **15**, 2379–2393 (2022).
[doi:10.1109/JSTARS.2022.3157740](https://doi.org/10.1109/JSTARS.2022.3157740)
51. Z. Wang, T. Li, R. Du, N. Yang, J. Ping, A high-efficiency lettuce quality detection system based on FPGA. *Comput. Electron. Agric.* **231**, 109978 (2025).
[doi:10.1016/j.compag.2025.109978](https://doi.org/10.1016/j.compag.2025.109978)
52. L. Zhang, C. Zhou, B. Liu, Y. Ding, H. J. Ahn, S. Chang, Y. Duan, M. T. Rahman, T. Xia, X. Chen, Z. Liu, X. Ni, Real-time machine learning-enhanced hyperspectro-polarimetric imaging via an encoding metasurface. *Sci. Adv.* **10**, eadp5192 (2024).
[doi:10.1126/sciadv.adp5192](https://doi.org/10.1126/sciadv.adp5192) [Medline](#)
53. Y. Xu, L. Lu, V. Saragadam, K. F. Kelly, A compressive hyperspectral video imaging system using a single-pixel detector. *Nat. Commun.* **15**, 1456 (2024). [doi:10.1038/s41467-024-45856-1](https://doi.org/10.1038/s41467-024-45856-1) [Medline](#)
54. P. Christen, D. J. Hand, N. Kirille, A review of the F-measure: Its history, properties, criticism, and alternatives. *ACM Comput. Surv.* **56**, 1–24 (2023). [doi:10.1145/3606367](https://doi.org/10.1145/3606367)
55. M. Kovar, M. Breštic, O. Sytar, V. Barek, P. Hauptvogel, M. Zivcak, Evaluation of hyperspectral reflectance parameters to assess the leaf water content in soybean. *Water* **11**, 443 (2019). [doi:10.3390/w11030443](https://doi.org/10.3390/w11030443)
56. D. Mullan, J. Pietragalla, “Leaf relative water content” in *Physiological Breeding II: A Field Guide to Wheat Phenotyping*, A. Pask, J. Pietragalla, D. Mullan, M. Reynolds, Eds. (CIMMYT, 2012), chap. 25, pp. 25–35.
57. M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, B. Scholkopf, Support vector machines. *IEEE Intell. Syst. Their Appl.* **13**, 18–28 (1998). [doi:10.1109/5254.708428](https://doi.org/10.1109/5254.708428)
58. V. K. Chauhan, K. Dahiya, A. Sharma, Problem formulations and solvers in linear SVM: A review. *Artif. Intell. Rev.* **52**, 803–855 (2019). [doi:10.1007/s10462-018-9614-6](https://doi.org/10.1007/s10462-018-9614-6)
59. J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence* (MIT Press, 1992).
60. D. Stoppa, A. Simoni, L. Gonzo, M. Gottardi, G. F. Dalla-Betta, Novel CMOS image sensor with a 132-dB dynamic range. *IEEE J. Solid-State Circuits* **37**, 1846–1852 (2002).
[doi:10.1109/JSSC.2002.804347](https://doi.org/10.1109/JSSC.2002.804347)
61. M. Vatteroni, D. Covi, A. Sartori, A linear-logarithmic CMOS pixel for high dynamic range behavior with fixed-pattern-noise correction and tunable responsivity. *Sensors. IEEE* **2008**, 930–933 (2008). [doi:10.1109/ICSENS.2008.4716593](https://doi.org/10.1109/ICSENS.2008.4716593)
62. S. W. Ing, G. C. Gerhard, A high gain silicon photodetector. *Proc. IEEE* **53**, 1714–1722 (2005). [doi:10.1109/PROC.1965.4344](https://doi.org/10.1109/PROC.1965.4344)
63. J. Heinonen, A. Haarahiltunen, M. Serue, D. Kriukova, V. Vähänissi, T. P. Pasanen, H. Savin, M. A. Juntunen, “Temperature dependency of responsivity and dark current of

nearly ideal black silicon photodiodes," *Proc. SPIE* **11682**, *Optical Components and Materials XVIII*, 1168207 (2021).

64. Y. Xue, Y. Han, Y. Wang, J. Li, J. Wang, Z. Zhang, X. Cai, H. K. Tsang, K. M. Lau, High-speed and low dark current silicon-waveguide-coupled III-V photodetectors selectively grown on SOI. *Optica* **9**, 1219–1226 (2022). [doi:10.1364/OPTICA.468129](https://doi.org/10.1364/OPTICA.468129)
65. F. Yang, Y. Hu, J. Ou, Q. Li, X. Xie, H. Han, C. Cai, S. Ruan, B. Xiang, Waveguide-Integrated MoS₂ Field-Effect Transistors on Thin-Film Lithium Niobate with High Responsivity and Ultra-Low Dark Current. *ACS Photonics* **12**, 2128–2136 (2025). [doi:10.1021/acspophotonics.4c02618](https://doi.org/10.1021/acspophotonics.4c02618)
66. S. E. Qian, Hyperspectral satellites, evolution, and development history. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **14**, 7032–7056 (2021). [doi:10.1109/JSTARS.2021.3090256](https://doi.org/10.1109/JSTARS.2021.3090256)
67. M. A. Folkman, J. Pearlman, L. B. Liao, P. J. Jarecke, "EO-1/Hyperion hyperspectral imager design, development, characterization, and calibration," *Proc. SPIE* **4151**, *Hyperspectral Remote Sensing of the Land and Atmosphere* (2001).
68. X. Chen, Y. Chen, Z. Ma, F. C. A. Fernandes, How is energy consumed in smartphone display applications? *Proceedings of the 14th Workshop on Mobile Computing Systems and Applications* 1–6 (2013).
69. L. Shi, X. Liu, L. Shi, H. T. Stinson, J. Rowlette, L. J. Kahl, C. R. Evans, C. Zheng, L. E. P. Dietrich, W. Min, Mid-infrared metabolic imaging with vibrational probes. *Nat. Methods* **17**, 844–851 (2020). [doi:10.1038/s41592-020-0883-z](https://doi.org/10.1038/s41592-020-0883-z) [Medline](#)
70. Y. Magariyama, S. Sugiyama, K. Muramoto, Y. Maekawa, I. Kawagishi, Y. Imae, S. Kudo, Very fast flagellar rotation. *Nature* **371**, 752 (1994). [doi:10.1038/371752b0](https://doi.org/10.1038/371752b0) [Medline](#)
71. H. Qin, T. Xu, P. Liu, J. Xu, J. Li, DMSSN: Distilled mixed spectral-spatial network for hyperspectral salient object detection. *IEEE Trans. Geosci. Remote Sens.* **62**, 1–18 (2024). [doi:10.1109/TGRS.2024.3379380](https://doi.org/10.1109/TGRS.2024.3379380)
72. W. Li, F. Yuan, H. Zhang, Z. Lv, B. Wu, Hyperspectral Object Detection Based on Spatial-Spectral Fusion and Visual Mamba. *Remote Sens.* **16**, 4482 (2024). [doi:10.3390/rs16234482](https://doi.org/10.3390/rs16234482)
73. D. Zhang, Y. Li, J. Geng, H. M. Kim, M. Ma, S. Wang, I. Kim, T. J. Wijaya, N. Higashitarumizu, I. K. M. R. Rahman, D. Urmossy, J. Bullock, A. Ozcan, A. Javey, Spectral kernel machines with electrically tunable photodetectors, Dryad (2025); <https://doi.org/10.5061/dryad.jh9w0vtpv>.