

# A Hyperspectral Imaging Guided Robotic Grasping System

Zheng Sun<sup>1</sup>, Zhipeng Dong<sup>1</sup>, Shixiong Wang<sup>1</sup>, Zhongyi Chu<sup>2</sup>, Member, IEEE,  
and Fei Chen <sup>†1</sup>, Senior Member, IEEE

**Abstract**—Hyperspectral imaging is an advanced technique for precisely identifying and analyzing materials or objects. However, its integration with robotic grasping systems has so far been explored due to the deployment complexities and prohibitive costs. Within this paper, we introduce a novel hyperspectral imaging-guided robotic grasping system. The system consists of PRISM (Polyhedral Reflective Imaging Scanning Mechanism) and the SpectralGrasp framework. PRISM is designed to enable high-precision, distortion-free hyperspectral imaging while simplifying system integration and costs. SpectralGrasp generates robotic grasping strategies by effectively leveraging both the spatial and spectral information from hyperspectral images. The proposed system demonstrates substantial improvements in both textile recognition compared to human performance and sorting success rate compared to RGB-based methods. Additionally, a series of comparative experiments further validates the effectiveness of our system. The study highlights the potential benefits of integrating hyperspectral imaging with robotic grasping systems, showcasing enhanced recognition and grasping capabilities in complex and dynamic environments. The project is available at: <https://zainzh.github.io/PRISM>.

**Index Terms**—Perception for Grasping and Manipulation, Software-Hardware Integration for Robot Systems, Grasping

## I. INTRODUCTION

HYPERSPECTRAL imaging offers unparalleled capabilities for precise material identification and analysis by capturing rich spectral information across dozens to hundreds of contiguous bands [1]. This technology has been widely applied in diverse fields, including remote sensing [2], biochemical analysis [3], and industrial sorting [4]. Recognizing its potential, researchers have recently begun exploring its integration into robotics to enhance material recognition and environmental perception [5]. Such advancements are critical for enabling robotic systems to perform complex grasping and decision-making manipulation tasks.

Despite the potential advantages, the adoption of hyperspectral imaging in robotics remains limited. First, existing linescan hyperspectral systems typically necessitate auxiliary

Manuscript received: January, 23, 2025; Revised March, 24, 2025; Accepted May, 26, 2025.

This paper was recommended for publication by Editor Júlia Borràs Sol upon evaluation of the Associate Editor and Reviewers' comments. This study was supported in part by the InnoHK initiative of the Innovation and Technology Commission of the Hong Kong Special Administrative Region Government via the Hong Kong Centre for Logistics Robotics.

<sup>1</sup> Zheng Sun, Zhipeng Dong, Shixiong Wang, and Fei Chen are with the Department of Mechanical and Automation Engineering, T-Stone Robotics Institute, The Chinese University of Hong Kong, Hong Kong SAR zhengsun@link.cuhk.edu.hk, zhipengdong@cuhk.edu.hk, sxwang@hklcr.hk, f.chen@ieee.org.

<sup>2</sup> Zhongyi Chu is with the School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100191, China chuzy@buaa.edu.

<sup>†</sup> Corresponding authors

Digital Object Identifier (DOI): see top of this page.

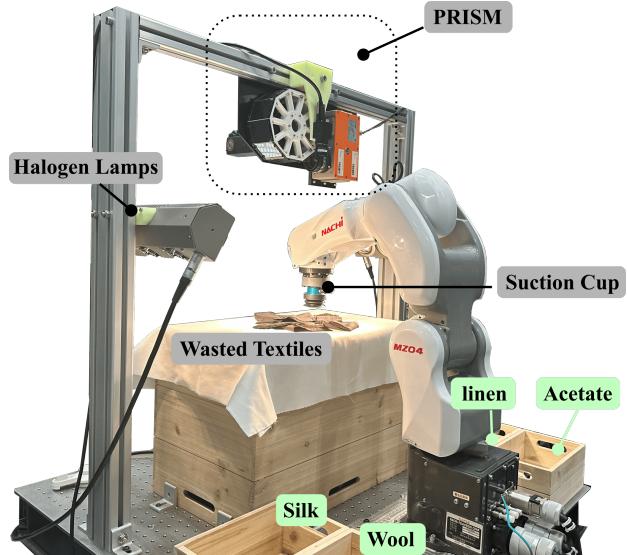


Fig. 1: Tabletop setup for the hyperspectral textile robotic grasping system. The setup includes a PRISM for hyperspectral image acquisition, halogen lamps, a Bernoulli's Principle suction cup gripper attached to a robotic arm, and a set of wasted textiles (linen, acetate, silk, wool) for sorting.

mechanisms, such as conveyor belts, rendering them more suitable for industrial-scale applications than for close-range robotic manipulation tasks [6], [7]. Meanwhile, commercially available snapshot hyperspectral cameras are often twice as expensive as linescan cameras and offer fewer spectral bands, thus constraining their practical deployment in robotic systems [8]. Furthermore, the inherent high-dimensionality of hyperspectral data introduces significant computational and storage challenges, hindering real-time application in dynamic environments [9]. Lastly, while conventional hyperspectral applications emphasize spectral analysis, the equally valuable spatial information is frequently overlooked. Effective exploitation of both spectral and spatial information could significantly enhance robotic perception and manipulation performance.

To address these challenges, this paper introduces a hyperspectral imaging-guided robotic grasping system that effectively integrates hyperspectral perception with robotic grasping tasks. The proposed system consists of two core components: PRISM, a compact, cost-effective hyperspectral imaging device providing high precision and distortion-free hyperspectral imaging, which was first introduced in [10]; and SpectralGrasp, a novel framework that exploits hyperspectral imagery to generate precise and effective grasping strategies. To the best of our knowledge, few studies have focused on hyperspectral imaging-guided robotic grasping. Our system effectively integrates hyperspectral data into close-range robotic grasping and manipulation tasks, as illustrated in Fig.

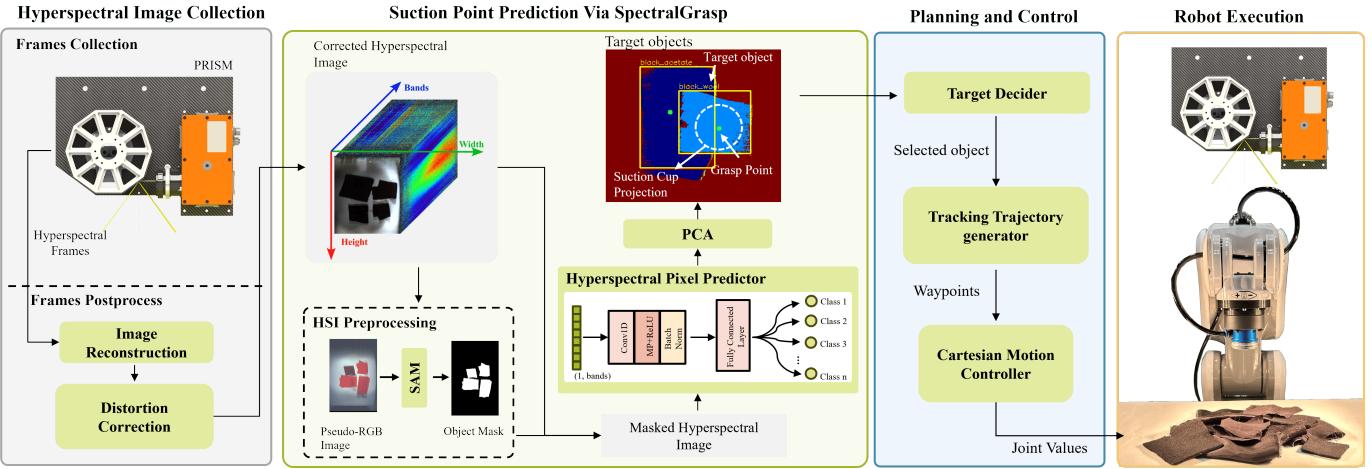


Fig. 2: System Overview: SpectralGrasp leverages both the spectral and spatial information from hyperspectral images captured by PRISM to generate suction points, enabling robots to complete textile sorting tasks. The process begins with PRISM capturing hyperspectral image frames, which are processed through image reconstruction and distortion correction to produce a corrected hyperspectral image. This corrected image with its mask serves as the input for SpectralGrasp. A pixel-level hyperspectral classifier is applied to generate pixel-level classification results. These results are then aggregated into object-level recognition outcomes using a Principal Component Analysis (PCA) algorithm. The system subsequently identifies the geometric centroids of target objects and generates a series of suction points for each target. A Target Decider module selects the target object, based on the task requirements, to generate the robotic execution trajectory. The trajectory is then sent to the Cartesian Motion Controller for precise sorting and execution.

1). The effectiveness of the proposed system is validated through textile sorting experiments, demonstrating superior recognition accuracy and sorting performance compared to human operators and conventional RGB-based methods. This study highlights the potential of hyperspectral imaging for enhancing robotic perception capabilities. This work makes the following contributions:

- 1) The design and implementation of PRISM, a hyperspectral imaging device achieving high-precision and distortion-free hyperspectral imaging with simplified integration into robotic systems.
- 2) The development of SpectralGrasp, a framework integrating spectral and spatial information from hyperspectral images to generate effective robotic grasping strategies.
- 3) Comprehensive validation through comparative experimental studies demonstrating improved recognition accuracy and sorting success rates relative to human performance and traditional RGB-based approaches.

## II. RELATED WORKS

### A. Robotics Sensing Technologies

Modern robotic manipulation heavily relies on perception, particularly in robotic grasp tasks. Early research primarily relied on RGB vision systems to capture an object's external shape and basic visual features such as color and texture [11]–[13]. As the need for more sophisticated perception increased, depth cameras were introduced to enhance 3D geometry detection, allowing robots to estimate objects' volume and spatial orientation more accurately [14]–[16]. To further improve perception, tactile sensors were integrated into robotic systems, allowing robots to perceive additional properties, such as force [17]–[19], temperature [20] and material stiffness [21]. These multimodel sensory inputs greatly improved the robustness of robotic grasping by offering multi-dimensional

information. Despite these advances, one critical limitation remains: the inability to distinguish materials with similar visual and physical properties. This gap highlights the need for more advanced sensing technologies that can incorporate additional information modalities, such as spectral data, to enable robots to achieve reliable material differentiation.

### B. Hyperspectral Technologies

Hyperspectral technology captures detailed spectral information across dozens to hundreds of contiguous bands, enabling precise material characterization and analysis [1]. This capability has been extensively used in fields, including biochemistry analysis, quality control within the pharmaceutical industry [3], [22], non-destructive food inspection [23], and industrial sorting [4]. Beyond its spectral resolution, hyperspectral images also provide spatial data, enabling pixel-by-pixel analysis, which is invaluable for applications requiring pixel-level analysis, such as environmental mapping and industrial sorting [2], [4]. The adoption of hyperspectral imaging in robotics has been slow due to deployment challenges. Traditional point spectrometers are effective for material analysis but lack spatial context, whereas linescan hyperspectral cameras require complex auxiliary mechanisms for operation. Meanwhile, snapshot hyperspectral cameras usually offer fewer spectral bands at significantly higher costs [8]. Furthermore, the high-dimensional nature of hyperspectral data demands significant computational resources, often impeding its use in real-time robotic applications [9].

### C. Applications of Hyperspectral Technologies in Robotics

Several studies have explored the integration of hyperspectral technology with robotics systems. One approach involves embedding spectrometers in robotic end-effectors for material identification. For example, spectrometer-equipped grippers

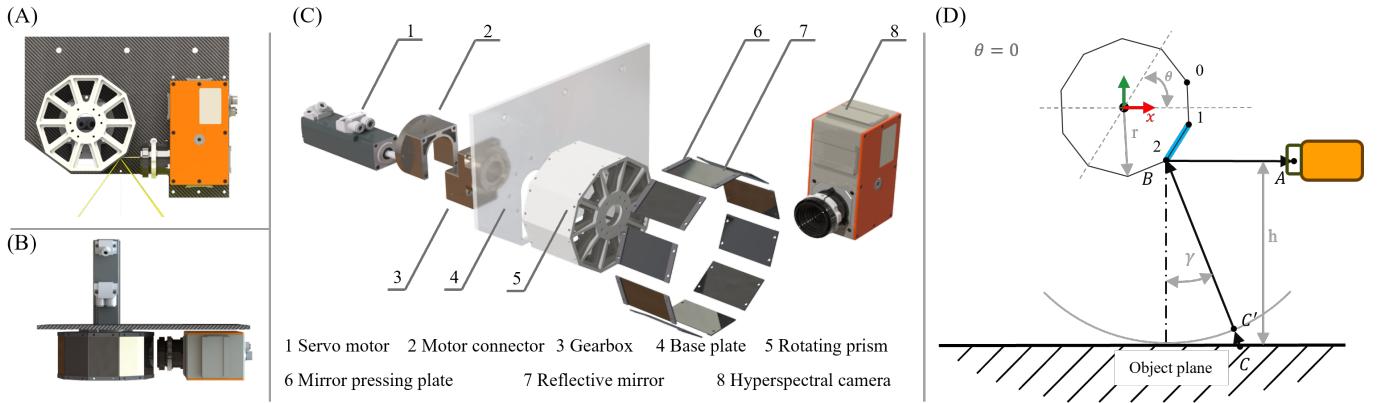


Fig. 3: CAD renderings of the mechanical design and the working mechanism of PRISM. (A) Fully assembled front view of PRISM, showing the two limited scanning positions. (B) Fully assembled top view. (C) Exploded view of PRISM, illustrating the individual components. (D) The working mechanism of PRISM, highlighting the scanning process. The angle  $\theta$  denotes the rotation angle of the servo motor, while the angle  $\gamma$  denotes the scanned angle. The circumradius of the rotating prism is  $r$ , and the position of the hyperspectral sensor, point  $A$ , is given by  $(x_A, \frac{\sqrt{2}r}{2})$ .

have been used to assess food quality, classify household materials and liquid classification in container [5], [24]–[28]. Some researchers have also attempted to enhance recognition capabilities in robotic grasping operations by integrating spectrometers into the gripper [29]. While these methods leverage spectral data, they often neglect the spatial information that could further enhance robotic perception.

Another line of research explores using robots to facilitate hyperspectral data acquisition. For instance, Wang *et al.* employed drones for remote sensing tasks [30], Azizi *et al.* used the precise movement of robots to carry objects under hyperspectral sensors for data collection [31]. And Hanson *et al.* developed an unmanned ground vehicle (UGV) equipped with hyperspectral imaging devices [32]; additionally, they also proposed a robotic pushbroom hyperspectral data acquisition system coupled with an effective algorithm for detecting occluded hyperspectral objects [33]. While these approaches leverage both the spectral and spatial information from hyperspectral images, they do not integrate this information into robotic manipulation or task execution.

To address these limitations, our proposed system leverages both spectral and spatial information to enable hyperspectral imaging-guided robotic grasping. By combining a custom-designed hyperspectral device, PRISM, with the SpectralGrasp framework, our work aims to bridge the gap between hyperspectral imaging and robotic manipulation, offering an effective solution for challenging material recognition and grasping tasks in close-range robotic scenarios.

### III. METHOD

The section details the design and implementation of the hyperspectral imaging device PRISM and the SpectralGrasp framework. These components integrate hyperspectral data into robotic systems by effectively leveraging spectral and spatial information to enhance material recognition and grasping performance. The proposed methodology highlights PRISM's innovative design for distortion-free hyperspectral imaging and demonstrates how SpectralGrasp generates robust grasping strategies from hyperspectral data, as illustrated in Fig. 2.

#### A. PRISM

*1) Design:* The PRISM (Polyhedral Reflective Imaging Scanning Mechanism) is a specialized hyperspectral imaging device optimized for robotic applications. PRISM comprises three primary components: a high-precision servo motor, a regular decagonal reflective prism, and a linescan hyperspectral camera, as illustrated in Fig. 3. PRISM simplifies the integration and operation of hyperspectral imaging systems by employing rotational reflection, thereby eliminating the need for complex conveyor mechanisms commonly required in traditional line-scan hyperspectral systems. During operation, the servo motor rotates the prism, sequentially reflecting the scanning line area onto the sensor of the hyperspectral camera via a reflective mirror. Compared to commercial snapshot hyperspectral cameras of equivalent spectral range, PRISM provides higher spectral resolution and precision at significantly lower costs.

The servo motor achieves control accuracy of 0.01 degrees, augmented by a high-precision encoder for real-time positional feedback. A gearbox with a reduction ratio of 10 enhances rotational accuracy and reduces torque requirements. Each facet of the decagonal reflective prism features high-reflectivity silvered mirrors, reflecting approximately 95% of incident light back onto the hyperspectral camera's sensor.

The number of sides of the reflective prism directly influences the PRISM's field of view (FOV). More sides result in a narrower scanning area but higher scanning precision. The relationship between the FOV and the number of the reflective prism sides  $n$  is defined by the following formula:

$$FOV = \frac{720^\circ}{n} \quad (1)$$

A decagonal prism configuration ensures uniform reflection and consistent imaging quality, providing an optimal balance between range and precision, yielding a practical FOV of 72°.

*2) Hyperspectral Image Reconstruction:* As shown in Fig. 3 (D), in the PRISM configuration, the scanned angle  $\gamma$  is a linear function of the motor angle  $\theta$ , meaning that for each motor position during the scanning cycle, there is

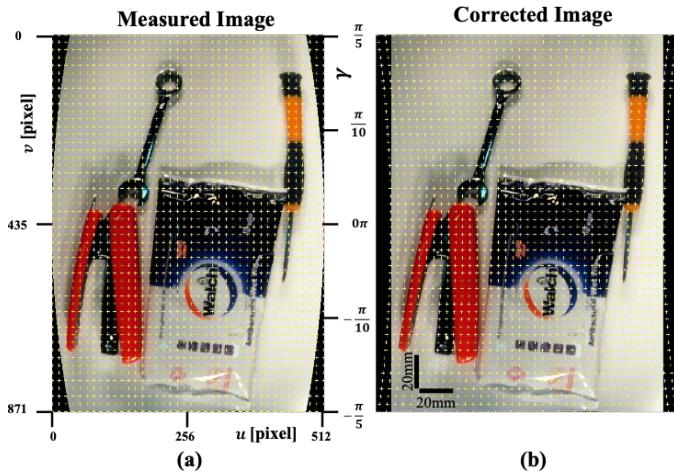


Fig. 4: Curvature distortion correction. (a) Original image exhibiting distortion with increased scanned angle  $\gamma$ . (b) Corrected image displaying uniform spatial mapping in both the  $x$  and  $y$  directions.

a corresponding unique scanning position. The relationship between the motor angle  $\theta$  and the scanned angle  $\gamma$  is:

$$\gamma = \frac{3\pi}{10} - 2\theta, \theta \in [\frac{\pi}{20}, \frac{\pi}{4}] \quad (2)$$

The operational range defined by  $\gamma$  ranges from  $\gamma \in [-\frac{\pi}{5}, \frac{\pi}{5}]$  facilitates uniform spatial resolution across hyperspectral images. Spatial resolution of PRISM hyperspectral images measures  $H \times W = 871 \times 512$  pixels,  $N$  representing the number of spectral bands. Each vertical direction position  $h$  corresponds to a unique motor angle  $\theta_h$ . During the imaging acquisition, each captured frame  $f \in \mathbb{R}^{1 \times W \times N}$  is recorded along with its corresponding motor angle  $\theta_h$ . These frames are then accurately mapped to their respective positions in the sample image based on the recorded angles, ultimately forming the complete hyperspectral image  $I \in \mathbb{R}^{H \times W \times N}$ . Typically, the motor is driven at 3 rpm. yielding a full-scan duration of approximately 2s.

3) *Distortion Correction*: Reflective mirrors introduce curvature distortion in the field of view (FOV). This distortion becomes increasingly pronounced as the scanned angle  $\gamma$  increases, visibly warping captured images, as illustrated in Fig. 4. Distortion in object width per scan line follows scaling factor  $k(\theta)$ , which is a nonlinear function of motor angle  $\theta$ :

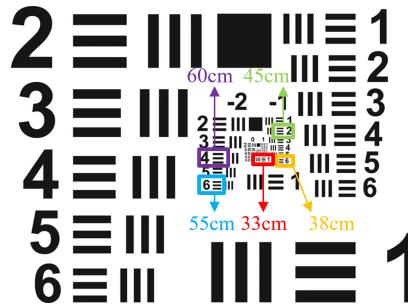
$$k(\theta) = \sqrt{1 + \frac{1}{\tan^2(\frac{\pi}{5} + 2\theta)}} \quad (3)$$

Each pixel coordinate  $(u, v)$  in the hyperspectral image  $I$  corresponds to specific spatial coordinates  $(x_d, y_d)$ , the relationship is described by:

$$y_d = h * \tan(\gamma) \quad (4)$$

$$x_d = u * \Delta x * k(\theta) \quad (5)$$

where  $\Delta x$  denotes the spatial resolution of each measured line in millimeters, determined by the camera lens properties and detection height  $h$ . To correct the curvature distortion, a transformation matrix  $M(\theta)$  is introduced incorporating the



	330	380	450	550	600
Working Height (mm)	330	380	450	550	600
Resolution (lp/mm)	4	3.56	2.24	1.78	1.41
Min Resolvable Size (mm)	0.13	0.14	0.22	0.28	0.35

Fig. 5: Spatial resolution test results for PRISM at various experimental heights. The resolution was measured using the 1951 USAF Resolution Test Chart (MIL-STD-150), containing reference line patterns with known dimensions. The diagram illustrates the achieved resolution in line pairs per millimeter (lp/mm) at heights ranging from 33 cm to 60 cm. The table below summarizes the spatial resolution performance and minimum resolvable size.

scaling factor  $k(\theta)$  to accurately map each pixel's spatial coordinates, resulting in the corrected hyperspectral image  $I^c$ .

$$I^c = M(\theta(u, v)) \cdot I(u, v) \quad (6)$$

The complete nonlinear transformation effectively removes distortion, as demonstrated in Fig. 4 (b). Although distortion has minimal impact on grasp estimation for non-detailed objects, the correction was utilized in the following experiments for optimal system performance.

### B. SpectralGrasp

1) *Overview*: This section introduces SpectralGrasp, a framework designed to utilize spectral and spatial information from hyperspectral images obtained by PRISM to generate robotic grasping strategies.

2) *Hyperspectral Imaging Preprocess*: As previously discussed, the substantial volume of hyperspectral data presents a significant challenge for real-time processing, given its complexity relative to standard RGB imagery. Identifying and isolating the region of interest (ROI) of target objects is therefore critical for enhancing real-time processing capabilities.

Initially, spatial object information is extracted from the corrected hyperspectral image  $I^c$ . Segmentation masks  $M \in \{0, 1\}^{H \times W}$  for each object are generated using the Segment Anything model [34] applied to pseudo-RGB images derived from  $I^c$ . These masks, combined with their respective hyperspectral data, produce masked hyperspectral images  $I^M$ , which proceed to further classification and localization steps. Redundant spectral information in hyperspectral images can negatively affect detection accuracy and processing speed; thus, a minimum noise fraction (MNF) method is applied to retain the most informative frequency bands [35]. The MNF method, implemented through the Python spectral library (spy) , typically compresses around thirty percent of spectral bands to optimize processing speed and recognition accuracy, effectively reducing computational load while preserving essential spectral information.

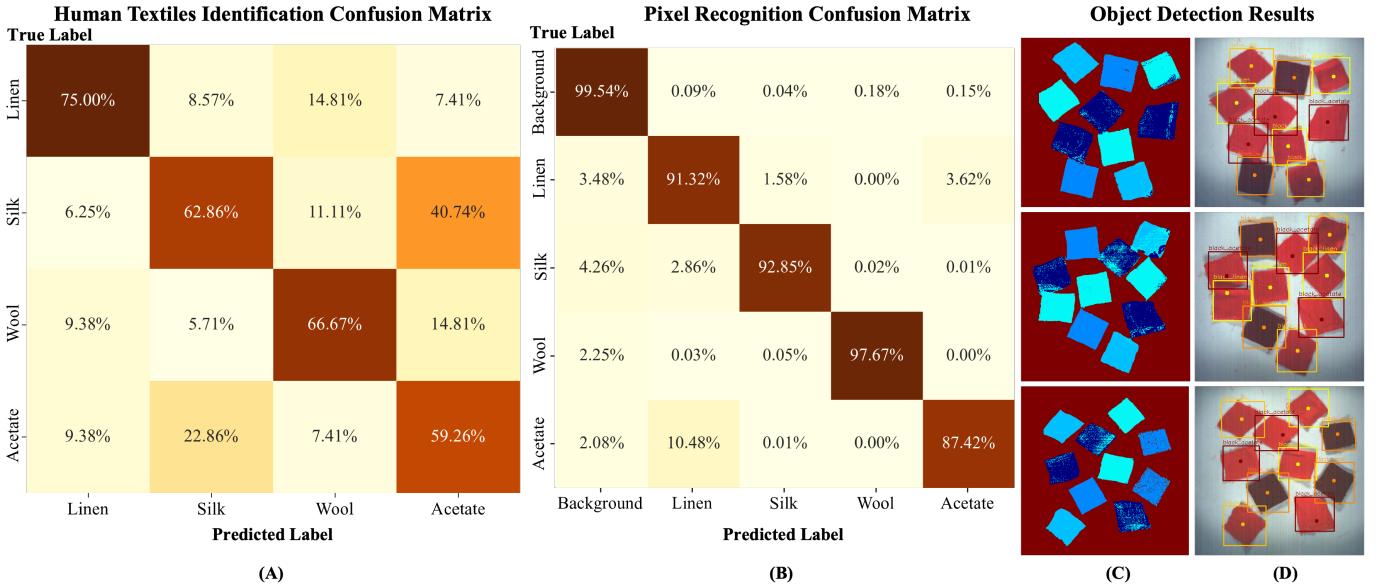


Fig. 6: Textile Recognition Experiment Results. Textile recognition experiment results. The comparison between human identification (A) and the system's pixel-level recognition results (B) highlights the significantly higher success rate of the proposed method. The object detection results demonstrate the effectiveness of our algorithm in processing segmentation maps (C) to extract semantic information (D), including object labels, bounding boxes, and suction points for following grasping execution.

**3) Object Detection and Grasp Point Prediction:** The objective is to classify each object and identify appropriate suction points. Traditional RGB-based detection models are inadequate due to hyperspectral image dimensionality. Thus, a specialized two-stage neural network for hyperspectral object detection is developed.

In the initial stage, pixel-level classification occurs within the masked hyperspectral image  $I^M$ . The spectral data of each pixel  $I^M(u, v)$  serve as inputs for the network, initially processed by one-dimensional convolution (Conv1d) layers to establish feature channels. Subsequently, one-dimensional max pooling (MaxPool1d) reduces spectral vector lengths for frequency-domain downsampling, followed by a ReLU activation function to introduce nonlinearity and batch normalization to standardize intermediate features. Iterating this convolution–pooling–normalization sequence reduces hyperspectral data dimensionality, resulting in compact feature representations, which are subsequently flattened and classified using a fully connected layer (nn.Linear).

The second stage aggregates pixel-level classifications into object-level classifications through Principal Component Analysis (PCA) clustering [36]. Geometric centroids of identified objects serve as candidate suction points. PCA's selection reflects its proven effectiveness for hyperspectral classification scenarios, though alternative methods such as Support Vector Machines (SVM) could also be implemented.

**4) Trajectory Planning:** A waypoint-based trajectory planner generates robot trajectories. Based on suction points and target bin locations identified from object classes, a sparse trajectory is initially established using predefined intermediate waypoints. This sparse trajectory undergoes refinement through Linear Quadratic Tracking [37], producing smooth Cartesian trajectories at 100 Hz for subsequent execution by the robot's Cartesian controller.

## IV. EXPERIMENT

A series of experiments was conducted to evaluate the performance of the PRISM device and the SpectralGrasp framework. These experiments assessed spatial imaging resolution, material recognition accuracy, and robotic sorting performance in practical scenarios.

### A. Spatial Resolution

PRISM employs a rotational scanning mechanism to eliminate the reliance on conveyor belts commonly required by linescan hyperspectral cameras. This design simplifies the deployment of hyperspectral imaging systems, enables the capture of static objects, and broadens the potential application scenario. Exploring the spatial resolution achievable through the rotational scanning function is essential to figure out the imaging performance of PRISM.

Spatial resolution, defined as the smallest distinguishable feature size, was quantified using the 1951 USAF Resolution Test Chart (MIL-STD-150), which contains standardized reference patterns. Tests at multiple heights ranging from 330 mm to 600 mm were performed, and the outcomes are illustrated as Fig. 5. The results demonstrate that PRISM can achieve a minimum resolvable size of 0.35mm at a working height of 600mm, which is sufficient for guiding close-range robotic perception and grasping tasks.

### B. Textile Recognition Experiment

PRISM's recognition performance and our algorithm's accuracy were evaluated by distinguishing various textile materials of identical color. Additionally, the effectiveness of PRISM was benchmarked against human performance, reflecting typical conditions in textile recycling factories, where fabric sorting often relies heavily on human visual inspection.

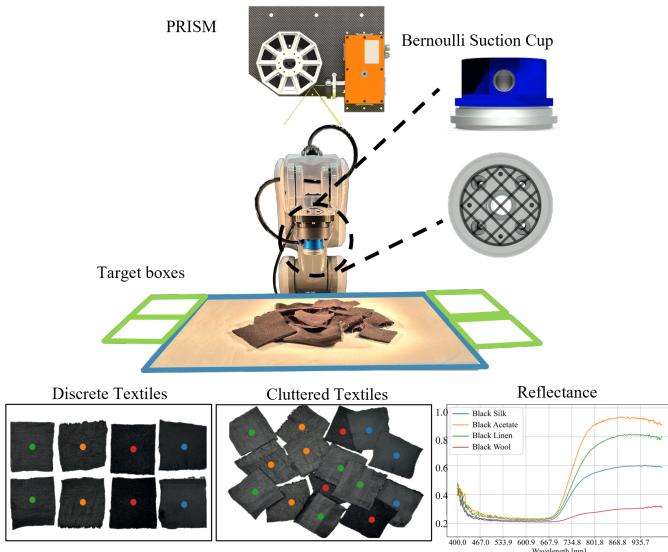


Fig. 7: Experiment setup for textile sorting experiment. Textiles are sorted from the object plane (surrounded by blue lines) to the corresponding target boxes (surrounded by green lines). Hyperspectral images are captured by the overhead PRISM system. We install a screen net to the Bernoulli suction cup's outlet to ensure effective suction. Performance is evaluated on two difficult levels black textiles, including silk, acetate, linen, and wool. Their spectrum curves are illustrated.

Four textile materials—linen, wool, acetate, and silk—each prepared in black, white, blue, and yellow colors, comprised the experimental dataset. Four trained participants independently performed textile identification tasks with mixed piles containing 16 distinct samples across four repetitions.

Results in Fig. 6 (A) showed an average human identification accuracy of 66%. In contrast, the pixel-level recognition classification results achieved accuracy ranging from 87% to frequently surpassing 91%, as illustrated in Fig. 6 (B). Object-level classification results further demonstrated robustness, as occasional pixel-level errors did not substantially influence object recognition accuracy or grasping point determination, as shown in Fig. 6 (C) and Fig. 6 (D). Overall, our algorithm robustly recognizes various materials and provides reliable grasping information, effectively leveraging hyperspectral data to guide robotic manipulation tasks.

### C. Hyperspectral Guided Sorting Experiment

A sorting experiment was conducted to verify the efficacy of PRISM and SpectralGrasp in real-world robotic sorting applications. The primary objective of this experiment is to demonstrate the efficacy of our hyperspectral imaging-guided robotic grasping system in accurately generating grasping points based on spectral-spatial information. Textile sorting is selected as the validation scenario because it effectively highlights the advantages of hyperspectral imaging in differentiating materials and object contours at pre-grasp stage, thereby enhancing grasping reliability.

*1) Setup:* For the hyperspectral guided sorting experiment, a NACHI MZ04 6-axis robotic arm was equipped with a Bernoulli suction cup gripper. This specialized suction cup enables the lifting of unsealed textile pieces without disturbing

the surrounding items. The suction cup is powered by a high-pressure airflow and controlled via a relay connected to a Modbus 485 module, ensuring consistent and reliable gripping performance. The PRISM is mounted above the workspace to capture global scene information. Four designated collection boxes are positioned around the workspace to receive the sorted textiles, as illustrated in Fig. 7. All experiments are conducted on a desktop computer running Windows 11, equipped with an Intel i7-13700KF CPU and an NVIDIA RTX 4090 GPU. The core codebase is implemented in C++ for computational efficiency, while additional Python scripts are utilized to simplify system interfacing.

---

#### Algorithm 1 Textiles Sorting Experiments

---

```

1: Initialize: robot.ready(), suction_cup.close()
2: Initialize: object_present ← true, target_bin
3: while object_present do
4:
    ▷ Stage 1: Capture and process hyperspectral image
5:      $\mathcal{I}^c \leftarrow \text{PRISM.capture().preprocess()}$ 
6:      $\mathcal{M} \leftarrow \text{SegmentObjects}(\mathcal{I}^c.rgb)$ 
7:     if  $\mathcal{M}$  is empty then
8:         object_present ← false
9:         break
10:    end if
11:
    ▷ Stage 2: SpectralGrasp
12:     $\mathcal{L}_{pixels} \leftarrow \text{PredictPixelLabels}(\mathcal{I}^c, \mathcal{M})$ 
13:    objects ← PCA( $\mathcal{L}_{pixels}$ )
14:
    ▷ Stage 3: Grasp pose calculation and execution
15:    for object ∈ objects do
16:         $\mathcal{G} \leftarrow \text{GetGraspPoses}(\text{object})$ 
17:        robot.move( $\mathcal{G}$ ), suction_cup.open()
18:        robot.move(target_bin[object])
19:        suction_cup.close()
20:    end for
21: end while

```

---

*2) Experimental Procedure:* The experiment aimed to recognize textiles and accurately sort them into the corresponding collection bins using the hyperspectral images obtained from PRISM and processed by the SpectralGrasp algorithm. The system's capability was assessed using four textile types: Linen, Silk, Wool, and Acetate, as shown in Fig. 7. The spectral reflectances of these textiles are also shown in Fig. 7. Two experimental conditions were designed to evaluate performance under varied complexities:

- 1) Discrete – Textiles are placed separately without overlap, allowing clear delineation of individual items.
- 2) Cluttered – Textiles are randomly arranged with overlaps, posing significant challenges in distinguishing boundaries.

Although previous comparisons with human identification demonstrated superior textile recognition capability of our proposed system, A comparative analysis against a state-of-the-art RGB-based detection method, YOLOv11 [13], was also performed. This approach allowed a direct comparison

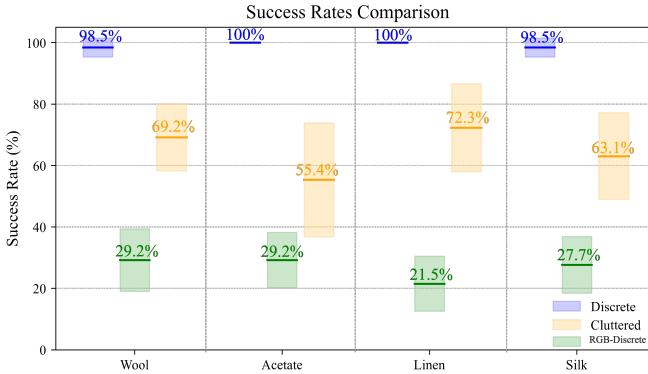


Fig. 8: Comparison of sorting success rates under three experimental conditions—Discrete, Cluttered and RGB guided Discrete—for four types of black textiles: Wool, Acetate, Linen, and Silk. The shaded regions indicate the mean  $\pm$  one standard deviation of the success rates, while the horizontal lines within these shaded areas mark the mean values.

of the sorting performance between RGB-guided sorting and our hyperspectral-guided approach, highlighting the value of spectral-spatial information.

52 individual samples were selected from four different types of textiles as the test set, with each sample typically measuring approximately  $5\text{cm} \times 5\text{cm}$ . In each trial, textiles were introduced into the sorting area according to the defined scenario, with minimal manual intervention. The complete experimental procedure is summarized in Algorithm 1.

3) *Results:* Performance results from five independent trials across the discrete hyperspectral, cluttered hyperspectral, and discrete RGB scenarios are summarized in Fig. 8. Under discrete conditions, hyperspectral-guided sorting achieved near-perfect performance. In contrast, RGB-guided sorting exhibited significantly lower effectiveness. The results strongly support hyperspectral imaging’s role in enhancing robotic grasping and sorting capabilities, particularly under scenarios challenging for traditional RGB methods. However, when transitioning from the relatively simple discrete scenario to the more complex cluttered scenario, sorting accuracy declined by 27% to 45% across the four textiles, with increased variability (standard deviation) in the results, demonstrating that pure hyperspectral sensing performance degrades in complex scenes. These findings do not imply that hyperspectral imaging universally outperforms RGB vision; rather, they underscore its value as a complementary sensing modality capable of significantly enhancing robotic perception and grasping capabilities in challenging environments.

#### D. Comparative Experiments

To the best of our knowledge, few studies have specifically addressed hyperspectral imaging-guided robotic grasping tasks, making it challenging to directly compare the grasp-point prediction accuracy of our proposed system with other existing methods. An indirect evaluation was therefore carried out by substituting the pixel-level classification module with alternative hyperspectral detection algorithms and measuring both recognition accuracy and computational efficiency. Benchmarks included the 3D-CNN approaches of Mei et al.

TABLE I: Comparative Experiment Results

Evaluation Parameter	Hyperspectral Recognition			
	Ours	Mei et al. [38]	Li et al. [39]	boulch et al. [40]
Training time (/h)	1.41	84	2.85	1.52
Inference time per image(/sec)	21.0	785	27.6	22.1
Accuracy (%)	98.02	96.6	98.52	<b>98.67</b>

[38] and Li et al. [39], as well as the 1D semi-supervised method of Boulch et al. [40]. Results in Table I show that the proposed network reduces training time to 1.41h and inference time to 21.0s per image—substantially lower than the 84h and 785s reported by Mei et al. and the 2.85h and 27.6s reported by Li et al.—making it far better suited for real-time grasping and manipulation scenarios. Accuracy remains competitive at 98.02%, differing by less than 0.7% from Li et al. (98.52%) and Boulch et al. (98.67%). Because most existing methods were developed for large-scale remote-sensing applications, where real-time performance is less critical, this balance of speed and accuracy underscores the suitability of the proposed classifier for object-level robotic tasks. The framework also allows other hyperspectral classifiers to be integrated as needed.

## V. CONCLUSION AND DISCUSSION

This study introduced a hyperspectral imaging-guided robotic grasping system, comprising PRISM and the SpectralGrasp framework, to enhance robotic perception and grasping accuracy. PRISM provides high-precision, distortion-free imaging with streamlined integration, while SpectralGrasp leverages spectral-spatial hyperspectral information to generate robust grasping strategies. Experiments involving textile recognition and sorting validated the system’s superior performance compared to human operators and RGB-based methods. Results emphasized the capability of hyperspectral imaging as a valuable perception modality in robotics. Future research directions include extending the proposed approach to more dynamic and complex manipulation scenarios. Additionally, integration of complementary sensing modalities, such as geometric or depth information, will be explored to further improve the robustness and versatility in diverse robotic grasping and manipulation tasks.

## REFERENCES

- V. Lodhi, D. Chakravarty, and P. Mitra, “Hyperspectral imaging system: Development aspects and recent trends,” *Sensing and Imaging*, vol. 20, pp. 1–24, 2019.
- J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, “Hyperspectral remote sensing data analysis and future challenges,” *IEEE Geoscience and remote sensing magazine*, vol. 1, no. 2, pp. 6–36, 2013.
- V. Silva-Perez, G. Molero, S. P. Serbin, A. G. Condon, M. P. Reynolds, R. T. Furbank, and J. R. Evans, “Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat,” *Journal of Experimental Botany*, vol. 69, no. 3, pp. 483–496, 2018.
- P. Tatzer, M. Wolf, and T. Panner, “Industrial application for inline material sorting using hyperspectral imaging in the nir range,” *Real-Time Imaging*, vol. 11, no. 2, pp. 99–107, 2005.
- Z. Erickson, E. Xing, B. Srirangam, S. Chernova, and C. C. Kemp, “Multimodal material classification for robots using spectroscopy and high resolution texture imaging,” in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 10452–10459.

- [6] G. ElMasry and D.-W. Sun, "Principles of hyperspectral imaging technology," in *Hyperspectral imaging for food quality analysis and control*. Elsevier, 2010, pp. 3–43.
- [7] Y. Zheng, J. Bai, J. Xu, X. Li, and Y. Zhang, "A discrimination model in waste plastics sorting using nir hyperspectral imaging system," *Waste Management*, vol. 72, pp. 87–98, 2018.
- [8] J.-B. Thomas, P.-J. Lapray, and S. Le Moan, "Trends in snapshot spectral imaging: Systems, processing, and quality," *Sensors*, vol. 25, no. 3, p. 675, 2025.
- [9] M. Makarenko, A. Burguete-Lopez, Q. Wang, F. Getman, S. Giancola, B. Ghanem, and A. Fratalocchi, "Real-time hyperspectral imaging in hardware via trained metasurface encoders," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 12692–12702.
- [10] Z. Sun, Z. Dong, S. Wang, and F. Chen, "Prism: A polyhedral reflective imaging scanning mechanism for hyperspectral applications," in *2024 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2024, pp. 1353–1358.
- [11] A. Bicchi and V. Kumar, "Robotic grasping and contact: A review," in *Proceedings 2000 ICRA. Millennium conference. IEEE international conference on robotics and automation. Symposia proceedings (Cat. No. 00CH37065)*, vol. 1. IEEE, 2000, pp. 348–353.
- [12] A. Saxena, J. Driemeyer, and A. Y. Ng, "Robotic grasping of novel objects using vision," *The International Journal of Robotics Research*, vol. 27, no. 2, pp. 157–173, 2008.
- [13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.
- [14] J. Redmon and A. Angelova, "Real-time grasp detection using convolutional neural networks," in *2015 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2015, pp. 1316–1322.
- [15] M. Gou, H.-S. Fang, Z. Zhu, S. Xu, C. Wang, and C. Lu, "Rgb matters: Learning 7-dof grasp poses on monocular rgbd images," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 13459–13466.
- [16] H.-S. Fang, C. Wang, H. Fang, M. Gou, J. Liu, H. Yan, W. Liu, Y. Xie, and C. Lu, "Anygrasp: Robust and efficient grasp perception in spatial and temporal domains," *IEEE Transactions on Robotics*, 2023.
- [17] T. Bhattacharjee, J. M. Rehg, and C. C. Kemp, "Haptic classification and recognition of objects using a tactile sensing forearm," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 4090–4097.
- [18] J. Lee, Z. Sun, Z. Dong, F. Chen, and H. S. Stuart, "Regrasping on printed circuit boards with the smart suction cup," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 6477–6483.
- [19] F. Chen, K. Sekiyama, P. Di, J. Huang, and T. Fukuda, "i-hand: An intelligent robotic hand for fast and accurate assembly in electronic manufacturing," in *2012 IEEE International Conference on Robotics and Automation*. IEEE, 2012, pp. 1976–1981.
- [20] S. Decherchi, P. Gastaldo, R. S. Dahiya, M. Valle, and R. Zunino, "Tactile-data classification of contact materials using computational intelligence," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 635–639, 2011.
- [21] J. Sinapov, V. Sukhoy, R. Sahai, and A. Stoytchev, "Vibrotactile recognition and categorization of surfaces by a humanoid robot," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 488–497, 2011.
- [22] T. W.-M. Fan and A. N. Lane, "Applications of nmr spectroscopy to systems biochemistry," *Progress in nuclear magnetic resonance spectroscopy*, vol. 92, pp. 18–53, 2016.
- [23] Y.-Z. Feng and D.-W. Sun, "Application of hyperspectral imaging in food safety inspection and control: a review," *Critical reviews in food science and nutrition*, vol. 52, no. 11, pp. 1039–1058, 2012.
- [24] V. Cortés, C. Blanes, J. Blasco, C. Ortiz, N. Aleixos, M. Mellado, S. Cubero, and P. Talens, "Integration of simultaneous tactile sensing and visible and near-infrared reflectance spectroscopy in a robot gripper for mango quality assessment," *Biosystems Engineering*, vol. 162, pp. 112–123, 2017.
- [25] Z. Erickson, N. Luskey, S. Chernova, and C. C. Kemp, "Classification of household materials via spectroscopy," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 700–707, 2019.
- [26] N. Hanson, W. Lewis, K. Puthuveetil, D. Furline, A. Padmanabha, T. Padir, and Z. Erickson, "Slurp! spectroscopy of liquids using robot pre-touch sensing," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 3786–3792.
- [27] S. Li, H. Yu, G. Pan, H. Tang, J. Zhang, L. Yé, X.-P. Zhang, and W. Ding, "M<sup>3</sup> tac: A multispectral multimodal visuotactile sensor with beyond-human sensory capabilities," *IEEE Transactions on Robotics*, 2024.
- [28] N. Hanson, G. Lvov, V. Rautela, S. Hibbard, E. Holand, C. DiMarzio, and T. Padir, "Prospect: Precision robot spectroscopy exploration and characterization tool," *arXiv preprint arXiv:2403.17232*, 2024.
- [29] N. Hanson, T. Kelestemur, D. Erdogmus, and T. Padir, "Pregrasp object material classification by a novel gripper design with integrated spectroscopy," *arXiv preprint arXiv:2207.00942*, 2022.
- [30] F. Wang, Q. Yi, J. Hu, L. Xie, X. Yao, T. Xu, and J. Zheng, "Combining spectral and textural information in uav hyperspectral images to estimate rice grain yield," *International Journal of Applied Earth Observation and Geoinformation*, vol. 102, p. 102397, 2021.
- [31] S. Azizi, E. Asadi, S. Howard, B. W. Muir, R. O'Shea, and A. Bab-Hadiashar, "Autonomous hyperspectral characterisation station: Robot aided measuring of polymer degradation," *IEEE Transactions on Automation Science and Engineering*, 2024.
- [32] N. Hanson, M. Shaham, D. Erdogmus, and T. Padir, "Vast: Visual and spectral terrain classification in unstructured multi-class environments," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 3956–3963.
- [33] N. Hanson, G. Lvov, and T. Padir, "Occluded object detection and exposure in cluttered environments with automated hyperspectral anomaly detection," *Frontiers in Robotics and AI*, vol. 9, p. 982131, 2022.
- [34] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo et al., "Segment anything," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 4015–4026.
- [35] J. B. Lee, A. S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 3, pp. 295–304, 1990.
- [36] C. Rodarmel and J. Shan, "Principal component analysis for hyperspectral image classification," *Surveying and Land Information Science*, vol. 62, no. 2, pp. 115–122, 2002.
- [37] H. Modares and F. L. Lewis, "Linear quadratic tracking control of partially-unknown continuous-time systems using reinforcement learning," *IEEE Transactions on Automatic control*, vol. 59, no. 11, pp. 3051–3056, 2014.
- [38] S. Mei, J. Ji, Y. Geng, Z. Zhang, X. Li, and Q. Du, "Unsupervised spatial-spectral feature learning by 3d convolutional autoencoder for hyperspectral classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6808–6820, 2019.
- [39] Y. Li, H. Zhang, and Q. Shen, "Spectral-spatial classification of hyperspectral imagery with 3d convolutional neural network," *Remote Sensing*, vol. 9, no. 1, p. 67, 2017.
- [40] A. Boulch, N. Audebert, and D. Dubucq, "Autoencodeurs pour la visualisation d'images hyperspectrales," in *Proc. 25th Colloque Gretsi*, 2017, pp. 1–4.