**Proceedings Paper:**

Alsari, A., Zhang, J., Farr, N.T.H. et al. (2026) Towards automated materials analysis: multi-scale spatio-spectral approach for registering secondary electron hyperspectral images. In: 2025 17th International Conference on Signal Processing Systems (ICSPS). 17th International Conference on Signal Processing Systems (ICSPS), 24-26 Oct 2025, Chengdu, China. Institute of Electrical and Electronics Engineers (IEEE). ISBN: 9798350392791.

https://doi.org/10.1109/ICSPS66615.2025.11347671

![University of Leeds logo] ![University of Sheffield logo] ![University of York logo]

# Towards Automated Materials Analysis: Multi-Scale Spatio-Spectral Approach for Registering Secondary Electron Hyperspectral Images

Ali Alsari
School of Electrical and
Electronic Engineering
*University of Sheffield, UK*
*ahalsari1@sheffield.ac.uk*

Jingqiong Zhang
School of Electrical and
Electronic Engineering
*University of Sheffield, UK*
*jingqiong.zhang@sheffield.ac.uk*

Nicholas T.H Farr
School of Chemical, Materials
and Biological Engineering
*University of Sheffield, UK*
*n.t.farr@sheffield.ac.uk*

Cornelia Rodenburg
School of Chemical, Materials
and Biological Engineering
*University of Sheffield, UK*
*c.rodenburg@sheffield.ac.uk*

Lyudmila Mihaylova
School of Electrical and
Electronic Engineering
*University of Sheffield, UK*
*l.s.mihaylova@sheffield.ac.uk*

*Abstract*—Secondary electron hyperspectral imaging (SEHI) is an emerging technique that allows experimental investigation of the chemical properties of material surfaces at the nanoscale. Despite advancements in automated image registration, registering SEHI images across different spatial scales remains a challenging task, and it is still typically performed manually in practice. Many image registration approaches often lead to inadequate registration in SEHI due to primarily focusing on spatial features, and do not capture the rich spectral information in SEHI. To fill in this gap, we propose a multi-scale spatio-spectral approach based on the scale-invariant feature transform (SIFT) to automate SEHI image registration. The main novelty of this work lies in the proposed multi-scale SIFT spatio-spectral descriptors that effectively integrate both the SIFT spatial descriptor and differentiate the spectral signal profile as a function of electron energy. A new spatio-spectral descriptor matching algorithm is designed, achieving accurate image registration, with an accuracy error of less than $0.44$ pixels. Results over real SEHI datasets show that the proposed approach performs better than other state-of-the-art registration methods. Two performance metrics are used for evaluation, including normalized cross correlation (NCC) for examining spatial alignment and cosine similarity regarding spectral alignment. The results show that our proposed approach achieves good accuracy, with an NCC of above $82\%$, and a cosine similarity of over $93\%$. This work offers a promising solution for automated multi-scale image registration, which is an essential step for SEHI analysis tasks for material surface chemical composition and accelerating materials discovery.

*Index Terms*—scanning electron microscopy, image registration, hyperspectral imaging, materials analysis and discovery, scale-invariant feature transform.

## I. INTRODUCTION

Scanning electron microscopy (SEM) is extensively used to understand composition and microstructure in micro- and nano-scale imaging, supporting a broad spectrum of industries and research areas, such as manufacturing, nanotechnology, and materials science [1]. A novel extension of SEM is secondary electron hyperspectral imaging (SEHI), which employs energy filtering within an SEM to produce a three-dimensional hyperspectral dataset [2]. SEHI data comprises high-resolution sequential images generated by secondary electrons emitted from the same sample area, with each image corresponding to a specific energy range that passes through a low-pass energy filter. SEHI has been demonstrated to be a valuable technique for advanced material applications such as surface chemical characterization, characterizing organic/inorganic material, and segmenting material surfaces [3], [4]. However, the advancements in imaging techniques from micro to nanoscale are still constrained by reliance on predominantly manual image registration. SEHI datasets across different magnifications rely on individual expertise for registration due to the absence of a reliable automated registration method, which is not only time-consuming but also inconsistent. To address this challenge, there is a need to develop an automated registration technique to enable precise, high-throughput characterization for both research and industrial purposes.

### A. Related Works

Image registration refers to the transformation of different image datasets into one coordinate system with matched imaging contents. Image registration can be mainly divided into two main parts: learning based registration and classical-based registration [5]. In recent years, learning based registration methods have integrated deep learning, which can be broadly categorized as end-to-end deep learning and feature extraction-based methods. Firstly, end-to-end methods can be divided into supervised and unsupervised learning. Supervised deep learning methods, such as [6]–[8], can directly predict transformation parameters but require extensive annotated training datasets comprising many image pairs with transformation parameters for network training. However, it is a costly task for SEHI images. Furthermore, training a deep convolutional

neural network (CNN) with limited data may not generalize well or may still need improvement. On the other hand, unsupervised learning techniques estimated transformation parameters by reducing the difference between image pairs [9]–[11]. These methods demonstrate superior performance in image registration and do not require extensive labeled datasets. However, they often struggle to register images with significant geometric distortions. Secondly, many researchers utilized deep neural networks as a feature extractor for image registration tasks [12]–[16]. Zhou et al. [16] proposed a multi-task unified deep CNN for remote sensing image registration and change detection. Quan et al. [15] introduced a deep wavelet learning network (DWNet) for feature learning. Zhao et al. [12] proposed an image registration algorithm based on a deep residual network (ResNet) and a traditional algorithm. SIFT and ResNet34 feature extraction were fused to improve the image registration. Ye et al. [14] fused traditional SIFT and convolutional neural network features to improve feature extraction and representation. Deep feature extractor methods have attained competitive registration accuracy by extracting high-level features in a nonlinear way. However, the problem of registration between images acquired from differing fields of view (FoV) cannot be handled by current methods.

Conventional registration methods can be divided into intensity-based registration and feature-based registration. Intensity-based registration worked directly with image intensity, such as the Fourier transform [17], phase correlation [18], and mutual information [19]. The primary constraint of this method is the lack of consideration for the spatial relationships across pixels, making it less appropriate for significant geometric changes. Feature-based image registration typically consists of four primary stages: detecting features, matching corresponding features, estimating the transformation model, and applying the transformation to align the images [20].

Numerous image registration methods have been proposed that leverage automated feature extraction to identify and align corresponding structures across images, such as the scale-invariant feature transformation (SIFT) [21], KAZE [22], Oriented Features from Accelerated Segment Test (FAST) and Rotated Binary Robust Independent Elementary Features BRIEF (ORB) [23], and many other algorithms. SIFT, which is invariant to scale, rotation, illumination, distortion, and viewpoint changes, is one of the most common feature detection and description algorithms. It can correctly locate key-points in noisy, cluttered, and occluded environments, which is the case of the SEHI images based on SEM for materials. SIFT offers superior accuracy and robustness compared to other methods, such as ORB and KAZE [24], leading to extensive application in the literature [24]–[28]. For example, a gradient definition and a robust matching strategy were introduced to improve the SIFT algorithm [27]. Besides, Zhang et al. [28] proposed a modified ratio of exponentially weighted averages (MROEWA) operator, and the traditional differences in the Gaussian (DoG) are replaced by Harris scale space to improve the SIFT algorithm for better detection of stable keypoints. This encourages us to inherit the powerful robustness of SIFT

to SEHI registration. However, the feature based methods are primarily designed to extract feature points from grayscale images, and when applied to hyperspectral imaging (HSI), they often operate on a single band from each image [29]–[31], or apply to images produced by dimension reduction, e.g, principal component analysis(PCA) [32]–[34].

Ordonez et al. [29] proposed HSI–KAZE method, which built a keypoint descriptor based on the spectral information and a nonlinear diffusion filter, to register hyperspectral remote sensing images. Liu et al. [35] proposed a remote sensing image registration framework based on ORB feature and optical flow theory for acousto-optic tunable filter (AOTF) spectrometer data. Zhang et al. [32] proposed a registration framework based on PCA for dimension reduction alongside global and local features of remote sensing HSI.

While the methods mentioned above can achieve good registration accuracy, they are not suitable for registering SEHI images. This is because the varying morphological characteristics of SEHI make it challenging to construct robust spatial and spectral descriptors, thus resulting in inaccurate feature matching, particularly when registering SEHI acquired under differing FoVs from diverse materials, and at a high spatial nano-resolution. In such cases, the registration process might be highly sensitive to small misalignment. Therefore, beyond the spatial features, the spectral information in SEHI should be well analyzed to enhance registration accuracy and material characterization. For instance, derivative spectral information with respect to the energy electron [36], [37] has been utilized effectively to construct a spectral descriptor in our proposed method, as shown in Fig. 1. To the best of our knowledge, the problem of registering SEHI data has not been previously investigated, and this paper is a preliminary study that introduces multi-scale spatio-spectral registration into SEHI images. Inspired by the prior research listed above, a multi-scale spatio-spectral registration approach is proposed for real-world SEHI applications towards automated materials characterization.

### B. Contributions

The main contributions of this work are as follows:

1. To the best of our knowledge, for the first time, we propose a multi-scale and spatio-spectral approach for SEHI image registration with different FoVs. Our approach adopts SIFT-based spatio-spectral descriptors, which allow fine-grained features in high-magnification images and coarse structures in low-magnification images to be well aligned.

2. Derivative-enhanced spectral data with respect to the energy electrons is first presented to construct the spectral descriptor, which effectively preserves material spectral signatures in SEHI data. This work also presents a new spatial-spectral descriptor matching algorithm that jointly integrates spatial and spectral information, achieving accurate SEHI image registration.

3. This study provides a comprehensive performance evaluation of the proposed approach on real-world cases, compared with other closely related registration methods. SEHI datasets

under four levels of magnification in SEHI imaging, covering FoV ranging from 30 μm to 100 μm, were utilized to validate the proposed approach and evaluate the performance.

4. The proposed approach can contribute to data-driven materials analysis and discovery by automating the process of nanoscale materials characterization for diverse applications, such as materials chemical composition analysis and materials discovery.

The rest of the paper is organized as follows: Section II introduces the proposed approach. Section III demonstrates the effectiveness of this approach for SEHI image registration. Finally, Section IV concludes this work.

## II. THE PROPOSED APPROACH

The proposed spatio-spectral SIFT registration approach has several stages which are detailed in Fig. 2 and consists of key steps summarised below.

### A. Smoothing

The 3D Gaussian filter used for smoothing the SEHI dataset in both spectral and spatial domains is expressed as:

$$G(x,y,\lambda,\sigma_1,\sigma_2) = \frac{1}{2\pi\sigma_1^2} \cdot \frac{1}{\sqrt{2\pi\sigma_2^2}} \cdot e^{-\left(\frac{x^2+y^2}{2\sigma_1^2} + \frac{\lambda^2}{2\sigma_2^2}\right)}, \quad (1)$$

where $x, y$, $\lambda$, $\sigma_1$, $\sigma_2$ are spatial coordinates, spectral coordinates, the standard deviation for spectral smoothing, and the standard deviation for spatial smoothing, respectively.

### B. Selected Band

The selected band image was obtained by selecting the band with the maximum entropy.

$$H_b = -\sum_{i=0}^{255} P_b(i) \cdot \log_2\left(P_b(i) + \varepsilon\right), \quad (2)$$

where $H_b$ is the entropy of band $b$, $P_b(i)$ is the probability of intensity level $i$ in band $b$ estimated from the normalized histogram , $\varepsilon$ is small constant (e.g., $10^{-8}$) added to avoid $\log(0)$. The maximum entropy band is selected because it captures the richest spatial and textural information and reduces the computational cost, making it well-suited for robust SIFT keypoint extraction. Bands with lower entropy tend to be more homogeneous or noisy, providing fewer distinctive features for matching. The use of maximum-entropy bands for initiating feature detection has also been reported in hyperspectral registration methods such as [25].

### C. Spatial Descriptor

Spatial feature extraction is achieved by the SIFT algorithm, and its primary steps are outlined next. The first is to define the local spatial feature based on the predefined band. Gaussian filters at different scales are used to blur the image, and then difference-of-Gaussian (DoG) images are generated by subtracting one blurred image from another. Stable keypoints $L(x,y,\sigma_2)$ across multiple scales are identified in this process

$$L(x,y,\sigma_2) = \left(\frac{1}{2\pi\sigma_2^2} e^{-\frac{x^2+y^2}{2\sigma_2^2}}\right) * I(x,y), \quad (3)$$

where $*$ indicates the convolution operation of $x$ and $y$ pixel coordinates and $I$ represents an image. The DoG $D(x,y,\sigma_2)$ is represented by:

$$D(x,y,\sigma_2) = (G(x,y,k\sigma_2) - G(x,y,\sigma_2)) * I(x,y) \quad (4)$$

Local maxima and minima in the DoG images are chosen as initial keypoints. Additionally, keypoints located in low-contrast regions or along edges are removed to improve stability and repeatability. Based on local image gradients, an orientation histogram is computed around the keypoint. The most dominant orientation is assigned to a keypoint to ensure rotational invariance. A region around the keypoint is divided into smaller subregions. Within each subregion, gradient magnitudes and orientations are computed to form histograms. These histograms are concatenated into a feature vector, which captures the local spatial structure and gradient patterns around the keypoint as presented in Fig. 3.

### D. Spectral Descriptor

The main steps for constructing the spectral descriptor are outlined next. mean spectral signature for each keypoint $k$ located at coordinates $(x_k, y_k)$, the region of interest (ROI) is defined as: $\text{ROI}_k = SEHI_{x,y,\lambda}$, where $x \in [x_k - r, x_k + r]$, $y \in [y_k - r, y_k + r]$, $r = t * m_s$, which represents the size of the region in pixels. A dynamically scaled size $r$ proportional to the median keypoint size $m_s$ from SIFT multiplied by tile size $t$, is implemented to ensure consistent and reliable ROI across varying magnifications. The mean spectral signature $mean_k$ for the ROI around each keypoint is computed as:

$$mean_k(\lambda) = \frac{1}{N} \sum_{x,y \in \text{ROI}_k} SEHI_{x,y,\lambda, } , \quad (5)$$

where $N$ is the total number of pixels in the region, $\lambda$ is the spectral band index, and $SEHI_{x,y,\lambda}$ represents the hyperspectral intensity at band $\lambda$ and spatial coordinates $(x,y)$. The spectral difference descriptor is computed using central-difference spectral differentiation across spectral bands with respect to energy electron [36], [37], which can mainly enhance material spectral signatures in SEHI data while simultaneously attenuating stochastic noise and compensating drift distortions. We incorporate these into our spectral descriptor to improve feature matching in downstream registration tasks across varied magnifications:

$$\Delta\text{mean}_k(\lambda) = \frac{1}{2}\left(\frac{\text{mean}_k(\lambda+1) - \text{mean}_k(\lambda)}{\Delta\text{eV}}\right.$$
$$\left. + \frac{\text{mean}_k(\lambda) - \text{mean}_k(\lambda-1)}{\Delta\text{eV}}\right), \quad (6)$$

where $\Delta eV$ is the energy step size between adjacent spectral bands, which is defined as: $\Delta eV = eV[i+1] - eV[i]$, where eV[i] is the energy electron value associated with the i-th band. The $eV$ is computed from mirror electrode voltage $(MV)$ according to the calibration equation: $eV = MV \cdot (-0.4446) + 6$, [36].
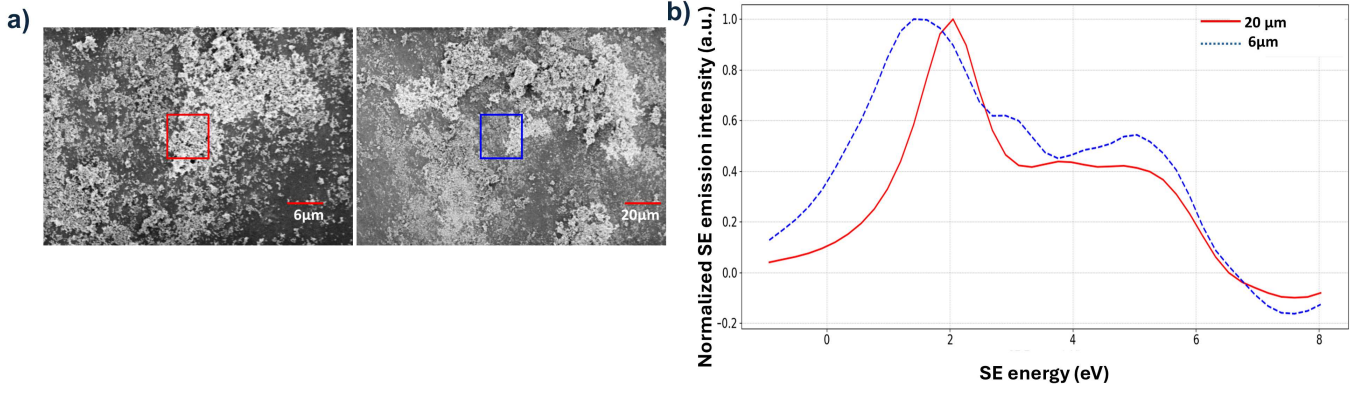
Fig. 1. a) SEHI data volume at 30μm (reference stack) and SEHI data volume at 100μm( target stack) b)Differentiate spectral information as a function of Secondary Electron (SE) energy Values (eV) derived from ROI of the SEHI data volumes.
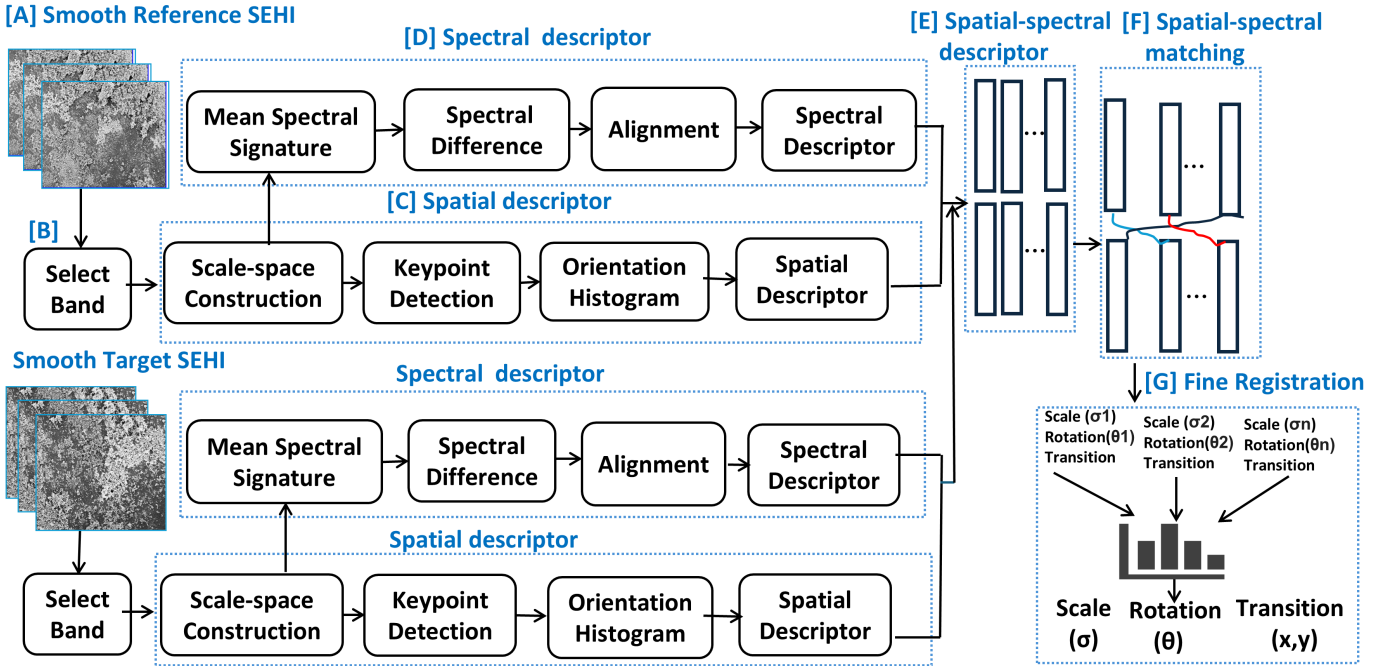


Fig. 2. Framework for multi-scale spatio-spectral registration of SEHI.

To align the spectral difference descriptors with the original spectral signature length, a zero padding is added to the start of the array:

$$D_k(\lambda) = \begin{cases} 0 & \text{if } \lambda = 0, \\ \Delta mean_k(\lambda - 1) & \text{if } \lambda > 0. \end{cases} \quad (7)$$

Finally, the descriptors are normalized for each keypoint:

$$\hat{D}_k(\lambda) = \frac{D_k(\lambda)}{\max\left(|D_k(\lambda)|\right) + \epsilon}. \quad (8)$$

### E. Spatial-Spectral Descriptor

We retrieved a spatial descriptor that includes 128 features by applying SIFT and a spectral descriptor of 48 elements through secondary electron spectra emission, with the methodologies described in sections C and D, respectively. The spatial-spectral descriptor is constructed by the concatenation of the spatial and spectral descriptors.

$$\mathbf{d} = \begin{bmatrix} \hat{\mathbf{x}}_s \\ \hat{\mathbf{x}}_e \end{bmatrix} \in \mathbb{R}^{176}, \quad (9)$$

where $\hat{\mathbf{x}}_s \in \mathbb{R}^{128}$ is the L2-normalized SIFT vector, $\hat{\mathbf{x}}_e \in \mathbb{R}^{48}$ is the normalized spectral vector. This fusion captures both the local image structure and the local secondary-electron energy signature in a single descriptor and is effective for downstream tasks, particularly SEHI descriptor matching.

### F. Spatial-Spectral Descriptor Matching

Descriptor matching follows two fundamental steps, as detailed in Algorithm 1. Initially, Given the set of $d$-dimensional spatial descriptors $D_s^{(1)}$ from the reference image and $D_s^{(2)}$
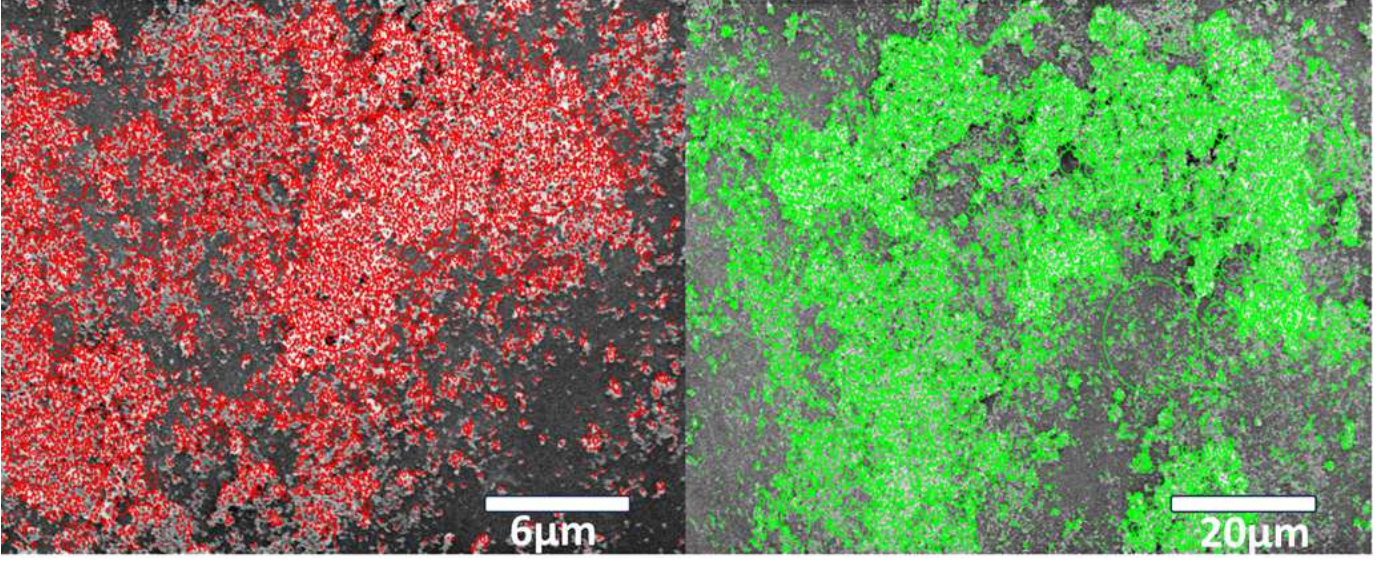
Fig. 3. Feature points of two SEHI images 30μm and 100μm calculated with SIFT.

from the target image, we first construct a $k$-d tree over $D_s^{(2)}$. For each descriptor $d_i \in D_s^{(1)}$, the tree is queried to retrieve its $k = 5$ nearest spatial neighbors $(j_1, d_1)$ and $(j_2, d_2),(j_k, d_k) \in D_s^{(2)}$, ranked by Euclidean distance. However, some descriptors may be missing in the target image, which causes erroneous matches. To address this, a Lowe's ratio test technique with a threshold of $\tau = 0.7$ is implemented to filter out incorrect matches. Only descriptor pairs satisfying $d_1 < \tau \cdot d_2$ are retained as valid spatial matches. Secondly, for a descriptor to be accepted as a good match, the cosine similarity between the spectral descriptors $D_p^{(1)}$ from the reference image and $D_p^{(2)}$ from the target image at the coordinate center of the corresponding spatial descriptor must exceed the spectral threshold $T_p = 0.90$. This step leverages spectral information to eliminate false matches effectively, ensuring higher accuracy in the final matching process, as shown in Figure 4.

### G. Fine Registration

In the fine registration stage, a full histogram-based search is carried out to register SEHI datasets with different magnifications. The approach determined a potential transformation for each pair of spatial-spectral descriptors matched. Subsequently, a selection is conducted based on all the rotation angles, transition, and scaling factors that have been obtained.

## III. EXPERIMENTS AND RESULTS

This section presents experiments to validate and examine the efficiency of the proposed approach. We evaluate our framework performance on four SEHI datasets as presented in the Table I namely $30\,\mu m, 40\,\mu m, 75\,\mu m$, and $100\,\mu m$ whose volume are $1000 \times 1500 \times 49$, and energy electron values (eV) start from -2 to 8. All experiments were conducted on a PC with an Intel 13th Gen Core-i7(24CPUs) @ 3.40 GHz, 32 GB RAM, and the algorithms were implemented in Python

---

**Algorithm 1:** Detailed procedure of the spatial–spectral matching algorithm

---

**Input:** $(D_s^{(1)}, D_p^{(1)}),(D_s^{(2)}, D_p^{(2)})$: spatial and spectral descriptors for SEHI dataset 1 and 2, $k$:number of neighbors, $\tau$:ratio-test threshold , $T_p$:spectral similarity threshold

**Output:** $M_{\text{good}}$: list of good matches passing both $\tau$ and $T_p$

$\text{KD2} := \text{BuildKDTree}(D_s^{(2)})$
$M_s := \emptyset$
**foreach** $i$ *in indices of* $D_s^{(1)}$ **do**
   $[(j_1, d_1), \ldots, (j_k, d_k)] := \text{KD2.KNN}(D_s^{(1)}[i], k)$
   **if** $d_1 < \tau\, d_2$ **then**
      $M_s := M_s \cup \{(i, j_1)\}$;

$M_{\text{good}} := \emptyset$
**foreach** $(m, n) \in M_s$ **do**
   **if** $\cos\!\big(D_p^{(1)}[m], D_p^{(2)}[n]\big) > T_p$ **then**
      $M_{\text{good}} := M_{\text{good}} \cup \{(m, n)\}$;

**return** $M_{\text{good}}$;

---

3.10 with OpenCV's SIFT 4.8.0, NumPy 1.24, and SciPy 1.11 libraries.

A cross-FoV registrations were performed, resulting in six distinct registration pairs, including (100μm, 75μm), (100μm, 40μm), (75μm, 40μm), (100μm, 30μm), (75μm, 30μm), and (40μm, 30μm). For each pair, we computed four performance metrics, namely registration error (pixels), NCC, cosine similarity, and spectral angle mapper (SAM).

We define the registration error as the mean Euclidean distance between matched feature points in the reference
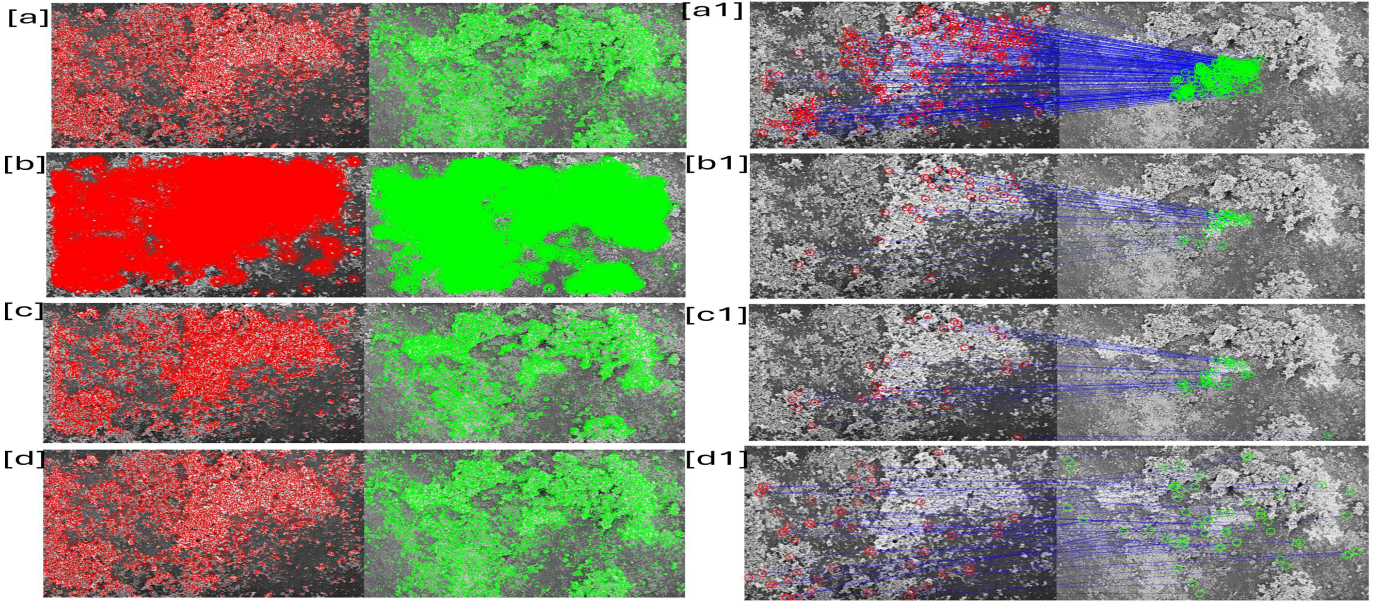
Fig. 4. [a–d]: Feature points on SEHI datasets (30μm, 100μm) calculated with our proposed method, HSI-ORB, HSI-KAZE, and HSI-SIFT-ResNet, respectively,[a1–d1]: Good matches from multi-scale spatio-spectral descriptor.
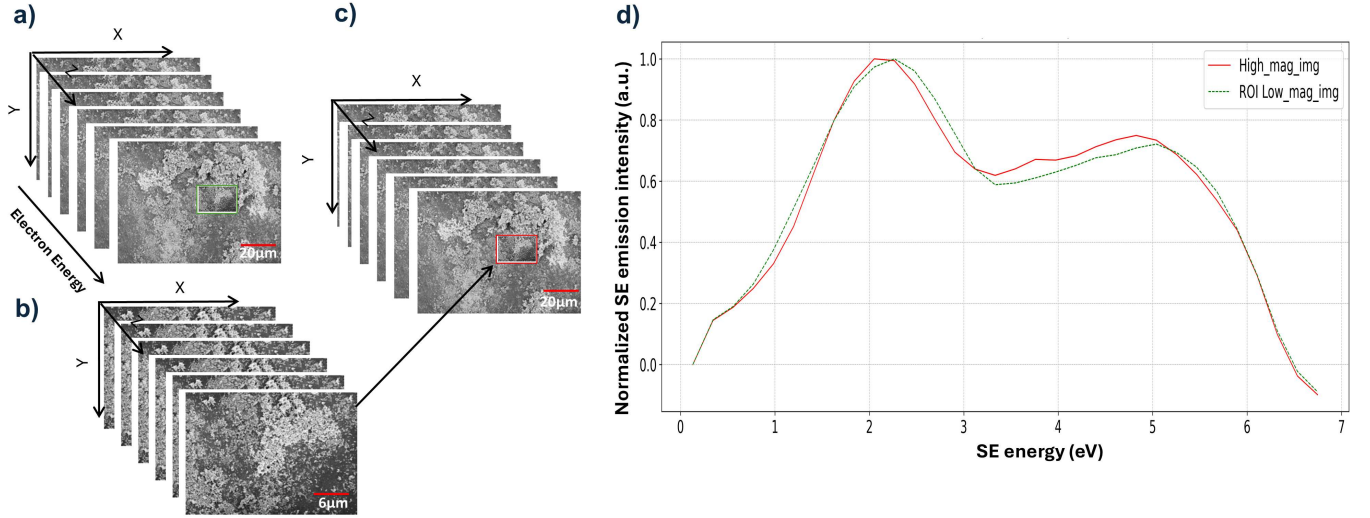


Fig. 5. a) SEHI data volume at 100μm (reference stack) b) SEHI data volume at 30μm( target stack) c) Registration target stack to align with reference stack d)Differentiate spectral information as a function of Secondary Electron (SE) energy Values (eV) derived from aligned regions of the SEHI data volumes.

image and their corresponding transformed points in the target image:

$$\text{Reg. Error} = \frac{1}{M} \sum_{i=1}^{M} \left\| \vec{r}_i - \vec{t}_i \right\|_2, \quad (10)$$

where $\vec{r}_i \in \mathbb{R}^2$ is the centre (pixel coordinates) of match feature point $i$ in the reference image, $\vec{t}_i \in \mathbb{R}^2$ is the corresponding matched feature point in the target image after applying the estimated transform, $M$ is the number of matches and $||.||_2$ denotes the Euclidean norm.

The NCC between the aligned images is defined as

$$\text{NCC}(R,T) = \frac{1}{N \, \sigma_R \, \sigma_T} \sum_{k=1}^{N} \left( r_k - \mu_R \right) \left( t_k - \mu_T \right), \quad (11)$$

where $R = \{r_k\}_{k=1}^{N}$ and $T = \{t_k\}_{k=1}^{N}$ are intensity values of the reference and aligned target images, respectively, $\mu_R, \mu_T$ are their means, and $\sigma_R, \sigma_T$ are their standard deviations.

For spectral similarity at the same spatial location, the cosine similarity and the SAM are identified as :

$$\cos \text{sim}(\mathbf{s}_r, \mathbf{s}_t) = \frac{\mathbf{s}_r^\top \mathbf{s}_t}{\|\mathbf{s}_r\|_2 \|\mathbf{s}_t\|_2}, \quad (12)$$

$$\text{SAM}(\mathbf{s}_r, \mathbf{s}_t) = \arccos\left(\frac{\mathbf{s}_r^\top \mathbf{s}_t}{\|\mathbf{s}_r\|_2 \|\mathbf{s}_t\|_2}\right), \qquad (13)$$

where $\mathbf{s}_r, \mathbf{s}_t \in \mathbb{R}^E$ are spectral vectors of reference and aligned target images with $E$ energy channels which are presented in Fig. 5. Each pair of reference and aligned target images is treated as an independent experiment, and the overall performance for each metric is computed by the mean

$$\bar{m} = \frac{1}{6}\sum_{j=1}^{6} m_j, \qquad (14)$$

where $m_j$ denotes the metric value for the $j$-th registration pair. Table II presents the results for the cross-FoV registration pairs alongside the mean. Across all heterogeneous FoV combinations, NCC for checking spatial alignment and cosine similarity for assessing spectra emissions alignment are greater than $82\%$ and $93\%$, respectively, and the registration error is less than $0.44$ pixels. The method achieved a mean registration error of 0.378 px, an NCC of 0.845, a cosine similarity of 0.973, and a SAM of 0.220 rad. The results demonstrate that the proposed approach effectively registers SEHI images across different FoVs and scales. The high values of NCC and cosine similarity measures, along with the low registration error, indicate that the framework is well-suited for registering SEHI images, particularly in complex materials and advanced manufacturing applications requiring precise alignment of SEHI images.

good match count, registration error (in pixels), NCC, and run-time on the SEHI dataset. The experimental results presented in these tables demonstrate that our proposed method significantly outperforms HSI-ORB, HSI-KAZE, and HSI-SIFT-ResNet in terms of registration precision and feature matching quality across both SEHI dataset settings ($100\,\mu$m, $30\,\mu$m) and ($100\,\mu$m, $40\,\mu$m). Specifically, the proposed method achieves the lowest registration error (0.41 px and 0.29 px) and the highest normalized cross-correlation (0.82 and 0.84), indicating superior alignment precision and structural consistency. While the runtime is slightly higher than HSI-KAZE, it remains comparable to HSI-ORB and significantly faster than HSI-SIFT-ResNet. Additionally, the proposed method achieves more matched keypoints—approximately 8.7× and 16.8× more than HSI-KAZE, 9× and 1.7× more than HSI-ORB, and 3.5× and 6× more than HSI-SIFT-ResNet, demonstrating its robustness and effectiveness for accurate registration in SEHI-based materials analysis.

TABLE III
COMPARISON OF DIFFERENT REGISTRATION ALGORITHMS ON SEHI DATASET (1000μM, 30μM)

| Method | Keyp1 | Keyp2 | Match | Err(px) | NCC | Time(s) |
|---|---|---|---|---|---|---|
| HSI-ORB | 20000 | 20000 | 29 | 0.57 | 0.77 | 24.66 |
| HSI-KAZE | 10357 | 10785 | 30 | 16.64 | 0.67 | **19.32** |
| SIFT-ResNet | 14094 | 28185 | 75 | 338.71 | 0.75 | 104.20 |
| Proposed | **14094** | **28185** | **262** | **0.41** | **0.82** | 27.25 |

TABLE I
DESCRIPTIONS OF FOUR SEHI DATASETS

| Dataset | Height (px) | Width (px) | Slices | HFW (μm) | Pixel size (nm) | eV Ranges (eV) |
|---|---|---|---|---|---|---|
| $100\,\mu$m | 1000 | 1500 | 49 | 100 | 66.7 | $-2.0 - 8.2$ |
| $75\,\mu$m | 1000 | 1500 | 49 | 75 | 50.0 | $-2.0 - 8.2$ |
| $40\,\mu$m | 1000 | 1500 | 49 | 40 | 26.7 | $-2.0 - 8.2$ |
| $30\,\mu$m | 1000 | 1500 | 49 | 30 | 20.0 | $-2.0 - 8.2$ |

TABLE IV
COMPARISON OF DIFFERENT REGISTRATION ALGORITHMS ON SEHI DATASET (100μM, 40μM)

| Method | Keyp1 | Keyp2 | Match | Err(px) | NCC | Time(s) |
|---|---|---|---|---|---|---|
| HSI-ORB | 20000 | 20000 | 194 | 0.62 | 0.77 | 26.46 |
| HSI-KAZE | 10911 | 10785 | 20 | 0.46 | 0.61 | **19.51** |
| SIFT-ResNet | 23063 | 28185 | 56 | 220.11 | 0.76 | 153.31 |
| Proposed | **23063** | **28185** | **335** | **0.29** | **0.84** | 25.30 |

TABLE II
METRICS FOR THE SIX CROSS-FoV REGISTRATIONS, AND THEIR MEAN

| Registration Pair | Reg. Error (px) | NCC | Cos* | SAM (rad) |
|---|---|---|---|---|
| (100μm, 75μm) | 0.26 | 0.86 | 0.97 | 0.26 |
| (100μm, 40μm) | 0.29 | 0.84 | 0.93 | 0.36 |
| (75μm, 40μm) | 0.44 | 0.85 | 0.99 | 0.16 |
| (100μm, 30μm) | 0.41 | 0.82 | 0.96 | 0.33 |
| (75μm, 30μm) | 0.44 | 0.83 | 0.99 | 0.16 |
| (40μm, 30μm) | 0.43 | 0.87 | 1.00 | 0.05 |
| **Mean** | **0.378** | **0.845** | **0.973** | **0.220** |

*: represents cosine similarity

We compared our method with HSI registration methods, including HSI-KAZE [29], HSI-ORB [35], and HSI-SIFT-ResNet [12], to verify the performance of the proposed method. To ensure a fair comparison, all baseline methods were configured with their optimal parameter settings on SEHI data. Tables III and IV present number of detected keypoints,

## IV. CONCLUSION

A multi-scale spatio-spectral registration approach of SEHI images is proposed, which can be applied for automating micro- and nano-structured materials analysis. This work designs SIFT-based spatio-spectral descriptors. Beyond the spatial descriptors, it includes spectral differentiation based on electron energy variations. The results on different real datasets demonstrate that the proposed approach can achieve accurate registration compared with other methods. This framework is expected to promote accurate image registration with the SEHI technique towards automated material chemical characterization down to the nanoscale, particularly for investigating complex carbon-based materials. It would be beneficial to deploy this framework across various material applications and integrate it with digital twins to enhance the advanced manufacturing process.

## REFERENCES

[1] C. Sun, S. Lux, E. Müller, M. Meffert, and D. Gerthsen, "Versatile application of a modern scanning electron microscope for materials characterization," *Journal of Materials Science*, vol. 55, pp. 13 824–13 835, 2020.

[2] N. Stehling, R. Masters, Y. Zhou, R. O'Connell, C. Holland, H. Zhang, and C. Rodenburg, "New perspectives on nano-engineering by secondary electron spectroscopy in the helium ion and scanning electron microscope," *MRS Communications*, vol. 8, no. 2, pp. 226–240, 2018.

[3] J. Zhang, J. Nohl, N. Farr, C. Rodenburg, and L. Mihaylova, "Unsupervised learning assisted secondary electron hyperspectral imaging for high-throughput cheminformatics analysis of materials," in *BIO Web of Conferences*, vol. 129. EDP Sciences, 2024, p. 10012.

[4] J. Nohl, J. Zhang, L. Mihaylova, S. Cussen, and C. Rodenburg, "An accessible secondary electron hyperspectral imaging approach to draw meaningful insights from scanning electron microscopy," in *BIO Web of Conferences*, vol. 129. EDP Sciences, 2024, p. 06003.

[5] H. O. Velesaca, G. Bastidas, M. Rouhani, and A. D. Sappa, "Multimodal image registration techniques: a comprehensive survey," *Multimedia Tools and Applications*, vol. 83, no. 23, pp. 63 919–63 947, 2024.

[6] H. R. Boveiri, R. Khayami, R. Javidan, and A. R. Mehdizadeh, "Medical image registration using deep neural networks: A comprehensive review," *Computers and Electrical Engineering*, vol. 87, p. 106767, 2020.

[7] X. Chen, A. Diaz-Pinto, N. Ravikumar, and A. F. Frangi, "Deep learning in medical image registration," *Progress in Biomedical Engineering*, vol. 3, no. 1, p. 012003, 2021.

[8] X. Jia, J. Bartlett, W. Chen, S. Song, T. Zhang, X. Cheng, W. Lu, Z. Qiu, and J. Duan, "Fourier-net: Fast image registration with band-limited deformation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, 2023, pp. 1015–1023.

[9] Y. Ye, T. Tang, B. Zhu, C. Yang, B. Li, and S. Hao, "A multiscale framework with unsupervised learning for remote sensing image registration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022.

[10] M. Papadomanolaki, S. Christodoulidis, K. Karantzalos, and M. Vakalopoulou, "Unsupervised multistep deformable registration of remote sensing imagery based on deep learning," *Remote Sensing*, vol. 13, no. 7, p. 1294, 2021.

[11] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "Voxelmorph: a learning framework for deformable medical image registration," *IEEE Transactions on Medical Imaging*, vol. 38, no. 8, pp. 1788–1800, 2019.

[12] X. Zhao, H. Li, P. Wang, and L. Jing, "An image registration method using deep residual network features for multisource high-resolution remote sensing images," *Remote Sensing*, vol. 13, no. 17, p. 3425, 2021.

[13] W. Lee, D. Sim, and S.-J. Oh, "A cnn-based high-accuracy registration for remote sensing images," *Remote Sensing*, vol. 13, no. 8, p. 1482, 2021.

[14] F. Ye, Y. Su, H. Xiao, X. Zhao, and W. Min, "Remote sensing image registration using convolutional neural network features," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 232–236, 2018.

[15] D. Quan, H. Wei, S. Wang, Y. Li, J. Chanussot, Y. Guo, B. Hou, and L. Jiao, "Efficient and robust: A cross-modal registration deep wavelet learning method for remote sensing images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 4739–4754, 2023.

[16] R. Zhou, D. Quan, S. Wang, C. Lv, X. Cao, J. Chanussot, Y. Li, and L. Jiao, "A unified deep learning network for remote sensing image registration and change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–16, 2023.

[17] Q. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of fourier-mellin transforms for image registration and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 12, pp. 1156–1168, Dec 1994.

[18] Z. Ye, J. Kang, J. Yao, W. Song, S. Liu, X. Luo, Y. Xu, and X. Tong, "Robust fine registration of multisensor remote sensing images based on enhanced subpixel phase correlation," *Sensors*, vol. 20, no. 15, p. 4338, 2020.

[19] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens, "Nonrigid image registration using conditional mutual information," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 19–29, Jan 2010.

[20] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[21] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1. IEEE, 2005, pp. 886–893.

[22] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "Kaze features," in *Proceedings of the European Conference on Computer Vision (ECCV)*. Florence, Italy: Springer, Berlin, Germany, October 2012, pp. 214–227.

[23] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to SIFT or SURF," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. IEEE, 2011, pp. 2564–2571.

[24] S. A. K. Tareen and Z. Saleem, "A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK," in *Proc. of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. IEEE, 2018, pp. 1–10.

[25] Ordóñez, Acción, F. Argüello, and D. B. Heras, "Hsi-mser: Hyperspectral image registration algorithm based on MSER and SIFT," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 12 061–12 072, 2021.

[26] W. Zhang and Y. Zhao, "An improved SIFT algorithm for registration between SAR and optical images," *Scientific Reports*, vol. 13, p. 6346, 2023.

[27] W. Ma, Z. Wen, Y. Wu, L. Jiao, M. Gong, Y. Zheng, and L. Liu, "Remote sensing image registration with modified SIFT and enhanced feature matching," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, pp. 3–7, 2017.

[28] H. Zhang, Y. Song, J. Hu, Y. Li, Y. Li, and G. Gao, "OS-PSO: A modified ratio of exponentially weighted averages-based optical and SAR image registration," *Sensors*, vol. 24, p. 5959, 2024.

[29] Ordóñez, F. Argüello, and D. B. Heras, "Alignment of hyperspectral images using kaze features," *Remote Sensing*, vol. 10, no. 5, p. 756, 2018.

[30] X. Zhang, C. Leng, Y. Hong, Z. Pei, I. Cheng, and A. Basu, "Multimodal remote sensing image registration methods and advancements: A survey," *Remote Sensing*, vol. 13, no. 24, p. 5128, 2021.

[31] S.-Y. Guan, T.-M. Wang, C. Meng, and J.-C. Wang, "A review of point feature based medical image registration," *Chinese Journal of Mechanical Engineering*, vol. 31, pp. 1–16, 2018.

[32] X. Zhang, S. Li, Z. Xing, B. Hu, and X. Zheng, "Automatic registration of remote sensing high-resolution hyperspectral images based on global and local features," *Remote Sensing*, vol. 17, no. 6, p. 1011, 2025.

[33] J. Lu and S.-i. Kamata, "Multiscale superpixel hgcn combining cnn for semi-supervised hyperspectral image classification," *Journal of Advances in Information Technology*, vol. 15, no. 8, 2024.

[34] Y. Ren, L. Liao, S. J. Maybank, Y. Zhang, and X. Liu, "Hyperspectral image spectral-spatial feature extraction via tensor principal component analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 9, pp. 1431–1435, 2017.

[35] H. Liu, B. Hu, X. Hou, T. Yu, Z. Zhang, L. Liu, J. Liu, and X. Wang, "Real-time registration of unmanned aerial vehicle hyperspectral remote sensing images using an acousto-optic tunable filter spectrometer," *Drones*, vol. 8, no. 7, p. 329, 2024.

[36] J. F. Nohl, N. T. Farr, M. R. Acocella, A. J. Knight, G. M. Hughes, J. Zhang, S. Robertson, S. Micklethwaite, S. Murphy, T. Motlová *et al.*, "Secondary electron hyperspectral imaging of carbons: new insights and good practice guide," *Advanced Science*, p. e01907, 2025.

[37] N. T. H. Farr, G. M. Hughes, and C. Rodenburg, "Monitoring carbon in electron and ion beam deposition within FIB-SEM," *Materials*, vol. 14, no. 11, p. 3034, 2021.