



OPEN

Evaluating segmentation methods for UAV-Based Spoil Pile Delineation

Sureka Thiruchittampalam^{1,2}, Bikram Pratap Banerjee³, Nancy F Glenn⁴ & Simit Raval¹✉

Mine waste dumps consist of individual, blob-like spoil piles, each with unique geological and geotechnical attributes that contribute to the overall stability of the dump. Manually characterising these individual spoil piles presents challenges due to issues of accessibility, safety risks, and time consumption. Analysis of remotely acquired images, through object-based classification, offers a promising solution for the effective identification and characterisation of individual spoil piles. However, object-based classification's effectiveness hinges on segmentation, an aspect often overlooked in spoil pile analysis. Therefore, this study aims to identify and compare various segmentation approaches to pave the way for image-based spoil characterisation. A comparative analysis is conducted between traditional segmentation approaches and those rooted in deep learning methodologies. Among the diverse segmentation approaches evaluated, the morphology-based deep learning segmentation approach, Segment Anything Model (SAM), exhibited superior performance compared to other approaches. This outcome underscores the effectiveness of incorporating morphological data and deep learning techniques for accurate and efficient segmentation of spoil pile. The findings of this study provide valuable insights into the optimisation of segmentation strategies, thereby contributing to the application of image-based monitoring of spoil piles and promoting the sustainable and hazard free management of mine dump environments.

Keywords Mean shift segmentation, Simple linear iterative clustering, Voronoi-based segmentation, StarDist segmentation, Segment anything model

The complexity inherent in the composition of coal spoils presents a formidable challenge in the endeavour to devise stable dump structures^{1,2}. Formulation of such structures is the initial phase of spoil pile characterisation, constituting the foundational elements of dumps. However, the on-site characterisation of individual spoil piles is encumbered by challenges, encompassing strenuous efforts, safety hazards, and susceptibility to subjective human biases. Employing a remote sensing-based approach may reduce the risks and limitations associated with on-site waste characterisation.

In recent advancements, the utilisation of a classification framework based on remote data gathering through miniaturised sensors attached to unmanned aerial vehicles (UAVs) has surfaced as an effective solution in the mining context^{3–5}. This methodology not only attests to an elevated standard of safety but also affords a cost-effective and temporally efficient alternative to conventional in-field characterisation protocols. Such an approach denotes a paradigm shift towards safer, more efficient, and objective spoil pile characterisation.

Image-based characterisation of spoil piles presents unique challenges due to the inherent variability in their topography and the dynamic effects of shadows and sun angles during image acquisition⁶. These factors can lead to misclassification errors, resulting in pixels belonging to the same class being assigned to different classes or pixels from different classes being grouped together⁷. Object-based classification approaches address these issues by grouping pixels into objects and assigning each object to a specific class based on its spectral and textural characteristics. This approach provides a more realistic and geometrically accurate representation of spatial features compared to traditional pixel-based classification methods. However, segmenting high-resolution remote sensing images of complex terrain like mining spoil piles remains a challenging task. The development of robust segmentation algorithms that can consistently perform well across all objects and locations in the image is hindered by the variability in acquisition parameters, such as illumination and sharpness. Additionally, segmentation quality is inherently subjective and heavily influenced by the user's definition of objects of interest

¹School of Minerals and Energy Resources Engineering, University of New South Wales, Sydney, NSW 2052, Australia. ²Department of Earth Resources Engineering, University of Moratuwa, Moratuwa 01400, Sri Lanka.

³School of Surveying and Built Environment, University of Southern Queensland, Toowoomba, QLD 4350, Australia.

⁴Department of Geosciences, Boise State University, Boise, ID, USA. ✉email: simit@unsw.edu.au

and the granularity of those objects⁸. Consequently, segmented images often exhibit two common flaws: over-segmentation, where objects of interest are fragmented into multiple pieces, and under-segmentation, where objects of interest are merged with their surroundings⁹. Segmentation is therefore a critical step in object-based image analysis, as the accuracy of classification directly depends on segmentation quality. Given the lack of studies specifically exploring segmentation algorithms for spoil pile delineation¹⁰, evaluating the suitability of different algorithms for user needs and assessing the quality of segmentation outputs is of significant value to the mining industry. This study evaluates the usability of existing segmentation algorithms for spoil pile delineation and compares their performance once the algorithms are optimised for this specific application.

Methods and materials

Methodological overview

The workflow employed for the comparative analysis of traditional segmentation methodologies, including mean shift segmentation, simple linear iterative clustering (SLIC), and Voronoi-based segmentation, alongside deep learning-based segmentation approaches such as Stardist segmentation and the Segment Anything Model (SAM), is given in Fig. 1. The colour-based segmentation techniques encompass mean shift segmentation and SLIC, while Voronoi-based segmentation, Stardist segmentation, and SAM are characterised as morphology-based segmentation methodologies.

The workflow commences with the creation of an orthomosaic and digital surface model (DSM), using geotagged RGB images acquired through UAV data acquisition. Subsequent to this, the generation of reference segments ensues, providing a foundational reference for the subsequent evaluation of segmentation outcomes derived from both traditional and deep learning-based algorithms. Within the framework of traditional segmentation methodologies, the determination of optimal parameters is achieved through a comparative assessment of Hoover metric scores, thereby facilitating the attainment of segmentation outcomes of optimal quality. Simultaneously, fine-tuning of deep learning-based approaches is conducted to achieve precise spoil pile segmentation. In the case of deep learning methodologies, the hillshade derived from DSM was used as the input. This integrative approach ensures a refined and accurate segmentation process, contributing to the enhancement of both traditional and deep learning-based segmentation methodologies within the context of complex terrain analysis.

Subsequently, the performance evaluation is executed by comparing the selected optimal traditional segmentation under each respective algorithm with the outcomes obtained from fine-tuned deep learning algorithms. This comparative analysis is conducted utilising the Hoover metric, providing a quantitative assessment of segmentation effectiveness. The systematic workflow presented in Fig. 1 ensures a structured evaluation of both traditional and deep learning-based segmentation methodologies in the context of UAV-derived products.

Study site and data acquisition

The study focused on a mining dump site employing paddock dumping in the Sydney basin, New South Wales, Australia (Fig. 2(e)). The data collection took place on 9th November 2022, using a DJI Matrice 300 RTK quadcopter equipped with a 45-megapixel RGB camera (Zemuse P1 model) featuring a full-frame CMOS sensor (35.9 mm × 24 mm) and a 35 mm lens with a 63.5° field of view. The camera was mounted on a three-axis stabilised gimbal. The UAV system was configured to capture optical image data at solar noon, flying at a height of 80 m, resulting in a ground sampling distance (GSD) of 1.22 cm. The image acquisition settings included automated capture mode with 80% forward and side overlaps.

Data processing and reference segment generation

A Structure from Motion (SfM) based photogrammetric stitching package, specifically Pix4D Mapper (Pix4D SA, Lausanne, Switzerland), was employed for processing raw images acquired during the UAV mission. The workflow associated with processing necessitated an alignment procedure to discern common features within overlapping images, thereby assisting in the estimation of the camera's spatial location and orientation for each individual image. This process was pivotal for ensuring the precise alignment of the images, a factor that is essential for the subsequent generation of accurate models. Subsequent to the alignment phase, Pix4Dmapper executed a camera calibration procedure to ascertain both the intrinsic and extrinsic parameters of the camera system. Intrinsic parameters, which encompass focal length, lens distortion, and the principal point, alongside extrinsic parameters, which detail the camera's spatial position and orientation in the scene, were evaluated. Ground Control Points (GCPs) were employed to enhance the precision of the camera positioning and orientation data, functioning as reference markers to improve the fidelity of the model. Pix4Dmapper utilised the GCPs to rectify any discrepancies in camera positioning and orientation, thereby enabling the georeferencing of the resultant orthomosaic and Digital Surface Model (DSM). Following the calibration of the camera and the integration of GCPs, Pix4Dmapper generated the final orthomosaic and DSM. The processing pipeline culminated in the generation of an orthomosaic (Fig. 2(a)), and a DSM (Fig. 2(b)) characterised by a spatial resolution of 5 cm. Furthermore, the hillshade representation (Fig. 2(c)) was derived by applying a sigmoidal stretch to the DSM with a specified strength level set at 3 and scaling factor of 2.

Reference segments were generated through a process of manual digitisation (Fig. 2(d)), focusing specifically on individual spoil piles within paddocks. This digitisation procedure was executed with reference to spatial data acquired through on-site field investigations and the visual interpretation of the orthomosaic. A total of two hundred thirty-two (232) reference segments were meticulously delineated within an area spanning 17,174 m².

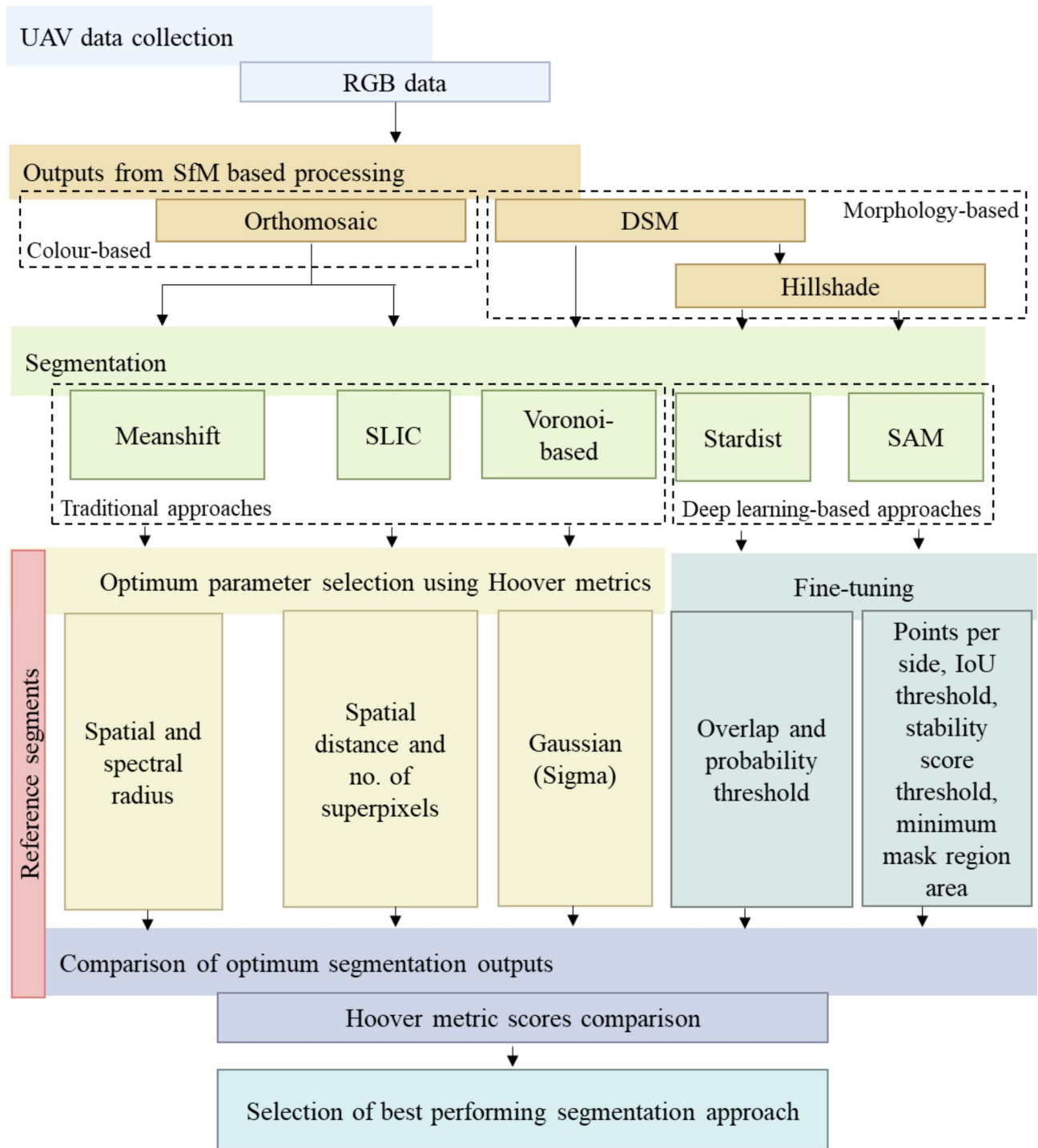


Fig. 1. Workflow employed for the comparative analysis of traditional and deep learning-based segmentation approaches in the context of coal spoil pile segmentation.

Segmentation

Object-based image classification comprises two sequential steps: orderly segmentation and subsequent classification. In the segmentation phase, pixels are systematically grouped into objects based on their homogeneity, which refers to the similarity of spectral values and spatial properties. This process involves the amalgamation of adjacent pixels that share common characteristics, facilitating the delineation of meaningful and coherent image objects. The objective is to create homogeneous segments that can be effectively employed for subsequent classification.

This study concentrated on investigating two distinct segmentation approaches: colour-based segmentation and morphological-based segmentation. Within the colour-based segmentation category, we explored two widely

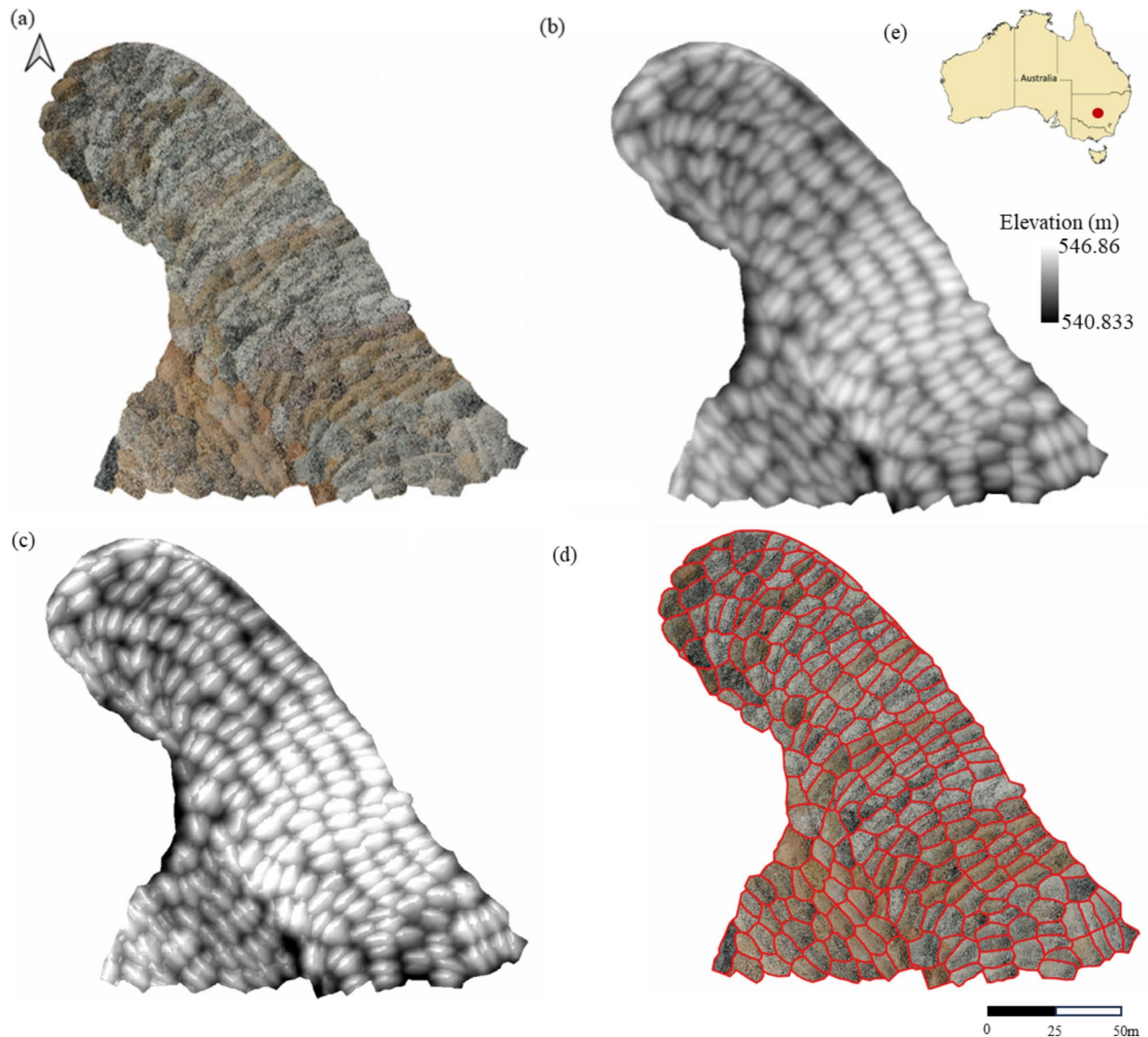


Fig. 2. (a) Orthomosaic and (b) digital surface model (DSM) and (c) hillshade representation of selected area of spoil piles, (d) reference segments and (e) location of selected area.

utilised iterative clustering algorithms, namely, the mean shift algorithm¹¹ and the SLIC algorithm¹². Within the morphological-based approach, we considered a traditional method known as Voronoi-based segmentation¹³. Additionally, two deep learning-based segmentation approaches, namely, StarDist segmentation¹⁴ and SAM¹⁵, were also explored in an attempt to assess their efficacy in the context of spoil pile delineation. A brief description of algorithms related to this work is provided in the below sections.

Mean shift segmentation

Mean shift is a non-parametric method characterised by an iterative process that relocates each data point to the average position of data points within its neighbourhood¹⁶. This algorithm identifies dense areas within a feature space, commonly referred to as clusters or regions of local maxima. Through an iterative clustering process, mean shift effectively groups the data points. The essential parameters for the mean shift algorithm segmentation include the spatial radius, range radius, and minimal region size.

In this study, various combinations of these parameters were employed, and the segmentation output generated with optimal parameter settings was chosen for comparative analysis with other segmentation algorithms.

Simple linear iterative clustering

Superpixel algorithms aim to generate perceptually meaningful clusters from individual pixels. The SLIC algorithm, a representative superpixel algorithm, is built upon the foundation of the k-means algorithm¹⁷. Unlike the k-means method, which operates across the entire colour space, the SLIC algorithm confines its

search region to a predefined area that encompasses the superpixel. This approach imparts certain advantages over the k-means method, including the capacity to control the compactness of superpixels. This is achieved by introducing a novel distance metric that considers both colour and spatial distances¹⁸.

The SLIC approach generates superpixels by clustering pixels based on colour similarity and proximity in the 'labxy' five-dimensional (5D) space¹⁷. 'lab' represents the pixel colour vector in the perceptually uniform CIELAB colour space, and 'xy' denotes the spatial position. While colour distance is constrained in CIELAB, spatial distance in the 'xy' plane depends on image dimensions, requiring normalisation for Euclidean distance in the 5D space. SLIC employs a distance measure accounting for superpixel size to ensure uniform cluster sizes and spatial extents, aiming for balanced colour and spatial influence in clustering for consistent superpixel creation.

In the context of this study, the optimisation of the SLIC algorithm involved considering parameters such as spatial distance and the number of superpixels. These parameters were systematically adjusted to achieve optimal segmentation results, aligning with the specific objectives and requirements of the study.

Voronoi-based segmentation

The Voronoi-based segmentation approach¹⁹ encompasses four steps: noise reduction, seed point detection, background seed point removal, and pile polygonisation based on Voronoi tessellation. The methodology initiates with the application of Gaussian blurring to the image for noise reduction, followed by the detection of local maxima. These steps contribute to the identification of seed points crucial for subsequent segmentation. Subsequently, Otsu's thresholding method²⁰ is employed to differentiate between the background and pile foregrounds. Background seed points are then eliminated through a binary operation involving the detected seed points and the background. The final step involves drawing polygonal areas corresponding to the piles based on Voronoi tessellation computed from the identified seed points. In this study, the sigma values of the Gaussian blur were systematically adjusted to optimise the algorithm for the precise detection of piles.

StarDist segmentation

The StarDist model²¹ (Fig. 3) employs a methodology wherein it anticipates a star-convex polygon for each pixel within the image. Precisely, it engages in the regression of distances from the pixel to the boundary of the corresponding object, considering a predetermined set of n radial directions characterised by equidistant angles²². This process is applicable exclusively to pixels within an object, excluding those associated with the background. Consequently, the StarDist model undertakes the discrete prediction of the likelihood of a pixel being part of an object, ensuring that polygons are derived solely from pixels with sufficiently elevated object probability. Following the generation of polygon candidates and their respective object probabilities, a non-maximum suppression (NMS) technique is implemented to yield the ultimate set of polygons, each representing a distinct object instance.

In contrast to a binary mask-based classification approach designating pixels as either object or background, the StarDist model adopts an alternative strategy. Specifically, the object probability of each pixel is defined as the normalised Euclidean distance to the nearest background pixel. This distinctive approach is strategically chosen to optimise the performance of the subsequent non-maximum suppression process. By ascribing object probabilities based on the Euclidean distance metric, NMS tends to prioritise polygons associated with pixels situated closer to the center of the object, a characteristic indicative of more accurate object representation.

Regarding the computation of distances for pixels within an object, the StarDist model calculates the Euclidean distances to the object boundary by traversing along each radial direction until encountering a pixel with a distinct object identity. This methodology ensures a comprehensive determination of distances for pixels constituting an object, facilitating subsequent analysis and predictions within the model's framework.

The StarDist algorithm is underpinned by the U-Net architecture. Specifically, it adopts the foundational structure of U-Net while incorporating a refinement in the form of an additional convolutional layer subsequent to the final U-Net feature layer. This strategic inclusion serves to mitigate potential conflicts between the two subsequent output layers, thereby enhancing the algorithm's efficacy. Notably, the ultimate layer in the architecture is responsible for generating the object probability. The implementation of StarDist is encapsulated within a Python package, complete with pre-trained models tailored for 2D applications. Of particular relevance is the utilisation of a model specifically trained on nuclei images, which aptly captures the morphological characteristics inherent in the hillshade dataset representing spoil piles central to the investigative focus of this work. The rationale behind employing this model lies in the discernible congruence between the morphological patterns observed in nuclei images and the hillshade depictions of the spoil piles.

This study utilised the hillshade representation of the DSM as input for StarDist segmentation, employing probability and overlap thresholds of 0.3 and 0.9 for optimal segmentation. Systematic adjustments of these parameters were implemented by varying the thresholds in different combinations. Subsequent to these adjustments, outcomes were subjected to visual assessment to ascertain the accuracy of the segmentation process.

Segment anything model (SAM) based segmentation

Meta AI has recently introduced a revolutionary foundation model called the SAM^{15,23}. Foundation models are deep neural networks trained using self-supervision and transfer learning on massive datasets, enabling them to adapt to various tasks and serve as the basis for diverse applications beyond those encountered during training. SAM represents a significant advancement in image segmentation methodologies. The SAM model's architecture comprises three core components (Fig. 4): an image encoder, a prompt encoder, and a mask decoder. The image encoder processes the input image, while the prompt encoder handles both sparse and dense prompts. The mask decoder generates a segmentation mask. This architecture allows SAM to predict multiple masks for the same instance.

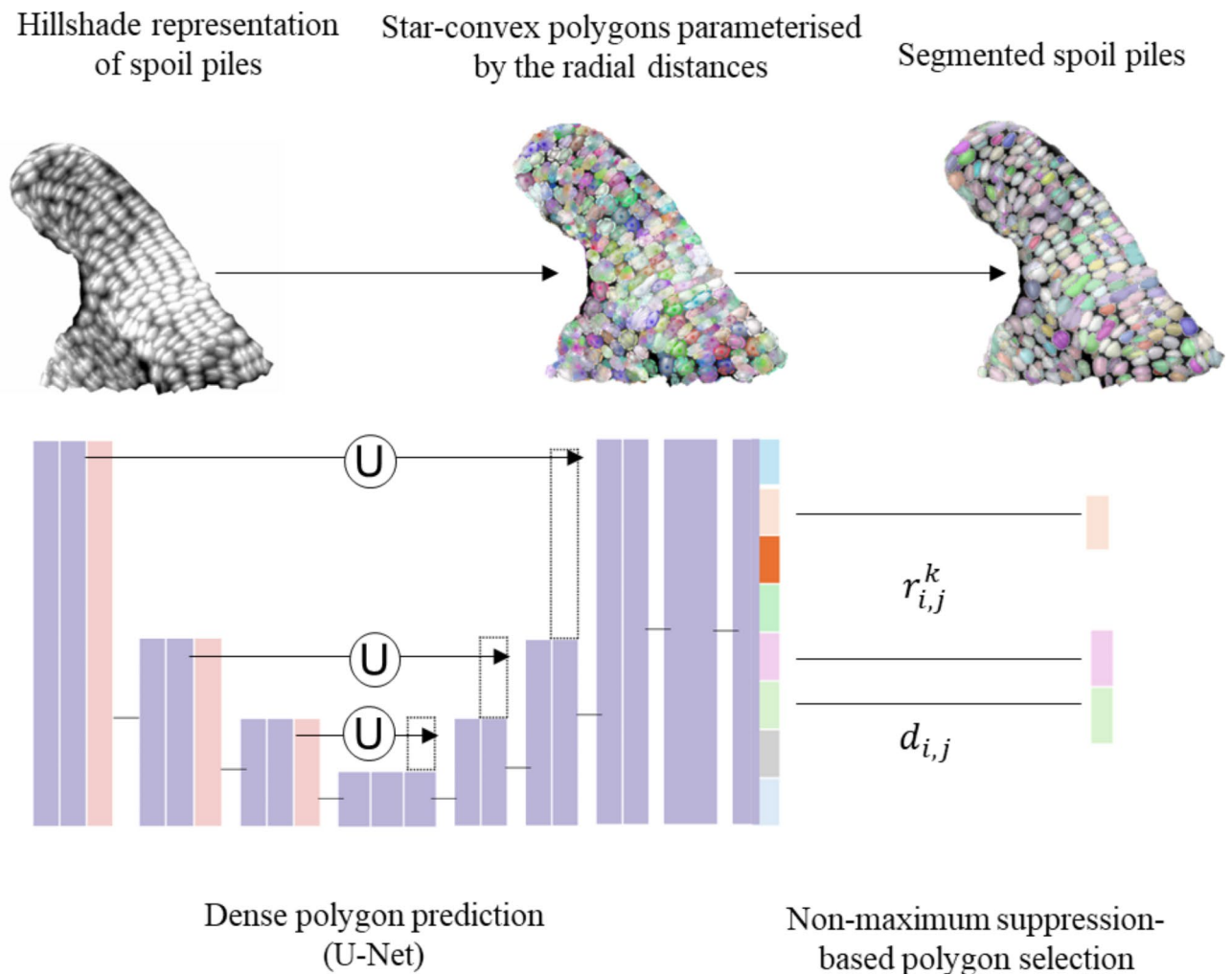


Fig. 3. Overview of StarDist. In this approach, each pixel $((i, j))$ in an image is characterised by two features: a star-convex polygon represented by radial distances from the pixel to the object's boundary ($r_{i,j}^k$), and the normalised Euclidean distance to the nearest background pixel ($d_{i,j}$), termed as “object probability”. These features are defined only for pixels within an object. The model predicts whether each pixel belongs to an object and considers polygon shape proposals only for pixels with high object probability. Using a U-Net backbone, the model predicts both radial distances and object probabilities. Finally, non-maximum suppression (NMS) is applied to all polygon candidates based on their object probabilities, favouring the polygon associated with the pixel near the object's center for accurate shape representation.

In this study, the pretrained SAM was utilised without conducting additional training from scratch. Instead of retraining the entire model, the study focused on fine-tuning specific parameters within the prompt encoder. This approach allowed the adaptation of the pretrained model to spoil pile delineation without the need for extensive computational resources and time. Fine-tuning involved adjusting the weights of the prompt encoder based on a task-specific dataset, which was carefully curated to reflect the nuances of spoil pile segmentation. In detail, SAM was applied to a hillshade representation of the spoil pile segmentation. A grid of points (200 points per side) was overlaid on the scene, and SAM generated all predicted masks, segmenting the entire scene at once. The study employed an intersection of union (IoU) threshold and stability score threshold of 0.9 and 0.9, respectively. Additionally, the number of crops, wherein an image crop denotes a smaller region extracted from a larger image to facilitate localised processing, was fixed at 1. The minimum mask region area, which was kept at 100 for optimal segmentation. These parameters were systematically adjusted, and the results were visually assessed to ensure accurate segmentation.

Performance evaluation of segmentations

This study investigated the optimal parameter settings for traditional segmentation techniques and evaluated their performance against alternative algorithms. The focus of the study was to determine object-level discrepancies between the reference segment and the segments generated by the algorithms. Hoover metrics were used for this purpose, as they evaluate boundary localisation accuracy and the effects of under- and over-segmentation, providing a more nuanced evaluation compared to pixel-level metrics. Hoover metrics²⁴, which

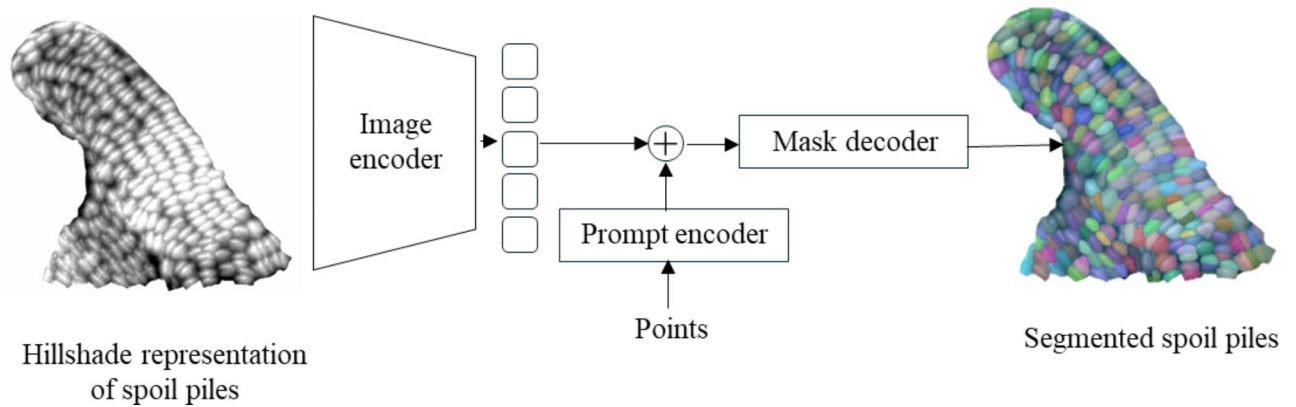


Fig. 4. Overview of the Segment Anything Model (SAM) architecture, which includes an image encoder, a prompt encoder, and a mask decoder. The image encoder utilises a Vision Transformer (ViT) to extract image embeddings. The prompt encoder integrates user interactions by encoding prompts into prompt embeddings. The mask decoder predicts segmentation masks by fusing the image and prompt embeddings using a lightweight transformer-based architecture.

quantifies the discrepancies between reference and machine-generated segmentations, were employed to assess the segmentation accuracy in terms of perfect segmentation, over-segmentation (a single object being partitioned into multiple objects), under-segmentation (inadequate partitioning of multiple objects), and missed detection (failure to detect an object). The segmentation evaluation was conducted using an overlapping threshold of 0.5. Four performance metrics, namely, correct detection score, over-segmentation score, under-segmentation score, and missed detection, were calculated and compared to comprehensively evaluate the segmentation performance of the algorithms. A higher score for each metric serves as a determinant in ascertaining whether the segmentation algorithm yields accurate segmentation, over-segmentation, under-segmentation, or missed segmentation.

Results and discussion

Optimal parameter selection for traditional segmentation approaches

The Hoover metrics scores for mean shift segmentation, SLIC segmentation, and Voronoi-based segmentation, along with their corresponding segmentation parameters, are presented in Fig. 5 (a), (b) and (c), respectively.

Mean shift segmentation exhibited improved segmentation performance when employing high spectral radii and low spatial radii at minimum region size of 10,000 pixels. The spectral radius serves as a parameter denoting the relative importance in discerning spoils based on their colour characteristics. A lower spectral radius is considered appropriate for distinguishing features characterised by closely aligned spectral attributes. Conversely, the spatial radius is a parameter reflecting the relative emphasis placed on segregating spoils according to their spatial characteristics. A lower spatial radius is deemed suitable for discriminating features that are both spatially confined and closely clustered. This finding (Fig. 5 (a)) can be attributed to the need for substantial smoothing to eliminate the fine granularity of individual piles, necessitating a low spatial radius as the optimal segmentation parameter. Conversely, the high spectral radius is attributable to the requirement for algorithms to differentiate between piles with similar spectral characteristics. While mean shift segmentation achieved the highest correct detection scores at a spectral radius of 20 DN and a spatial radius of 5 pixels, it is noteworthy that over-segmentation was also high (0.319) with these parameter settings. This indicates that even though the segmentation using algorithm does not precisely match the reference segmentation, it can still contribute to accurate object-based classification. This is because over-segmented polygons also fall within the geometry of reference segments and share their material properties. Therefore, both correctly detected and over-segmented piles, culminating in a Hoover score of 0.614, contribute to accurate classification.

The SLIC algorithm's performance (Fig. 5 (b)) exhibits a peak correct detection score of 0.365 at spatial distances of 30 and 550 superpixels. In the context of large superpixels, spatial distances take precedence over colour proximity, thereby affording greater relative significance to spatial proximity than to colour. This phenomenon results in the generation of compact superpixels that exhibit limited adherence to image boundaries. Conversely, for smaller superpixels, colour proximity supersedes spatial distances, assigning greater relative importance to colour proximity than to spatial proximity¹⁷. In this specific instance, a spatial distance of 30 and 550 superpixels has been identified as an appropriate parameter, facilitating the delineation of more discernible spoil piles. However, the substantial heterogeneity within the piles leads to over-segmentation of most piles, yielding an oversegmentation score of 0.375. This over-segmentation stems from SLIC's pixel clustering approach, which utilises both spectral similarity and proximity within piles. Consequently, the aggregate contribution of correctly and oversegmented piles to accurate classification amounts to 0.74 in Hoover scores.

The Voronoi-based segmentation algorithm's performance (Fig. 5 (c)) is maximised when the sigma value of Gaussian blurring is set to 12. This optimal performance is attributed to the balanced noise removal achieved at this sigma value. Sigma values lower than 12 fail to adequately eliminate noise, hindering the identification of individual pile maxima. Conversely, sigma values higher than 12 blur the images to an extent that distinct piles

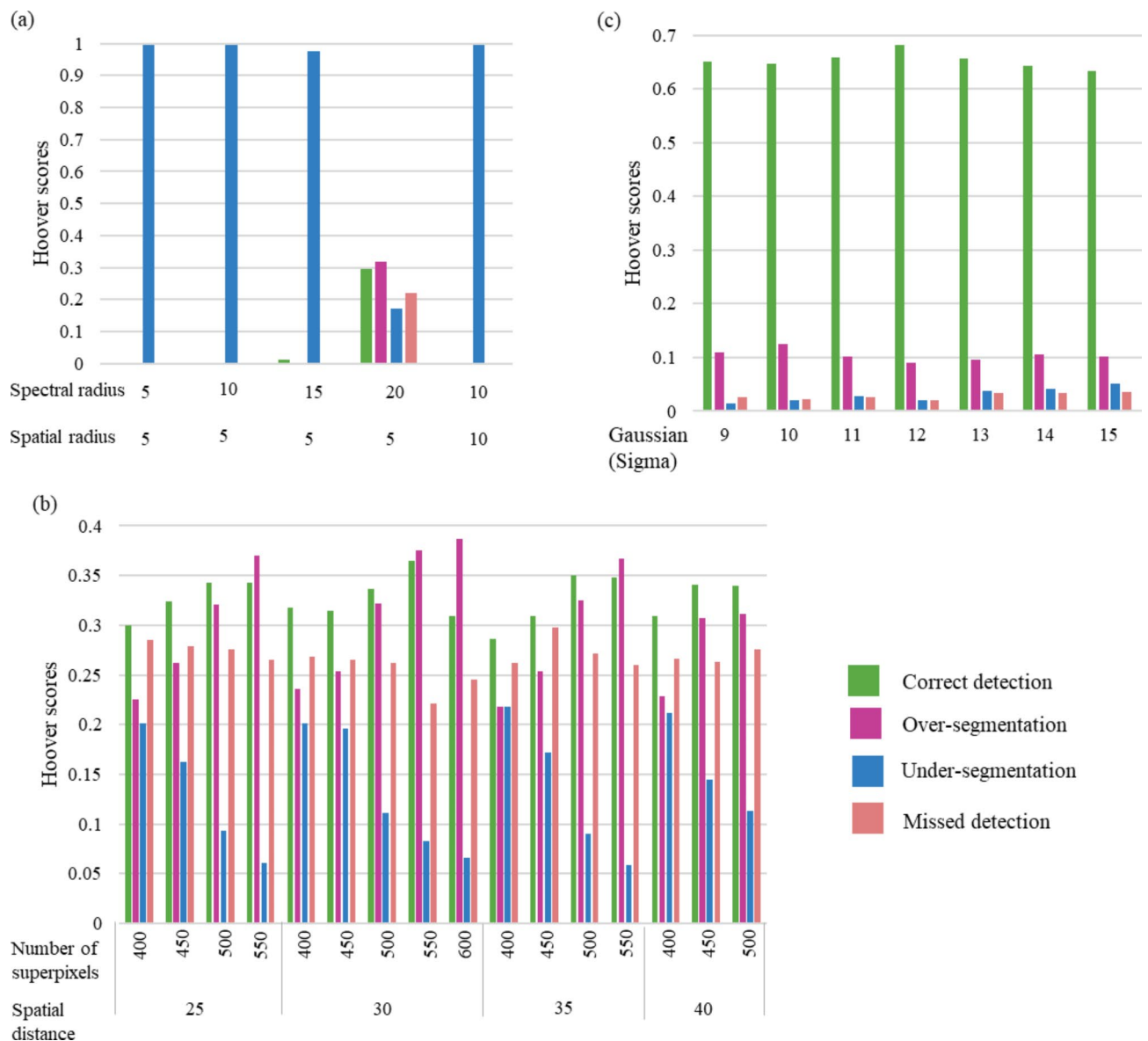


Fig. 5. Hoover metric scores corresponding to parameters in (a) mean shift segmentation, (b) Simple Linear Iterative Clustering (SLIC) and (c) Voronoi-based segmentation.

are no longer distinguishable for local maxima detection. Notably, both correctly and over-segmented piles at sigma value of 12 exhibit a Hoover score of 0.77.

Optimum segmentation using deep learning-based approaches

StarDist identifies individual piles by predicting the distances to object boundaries along a fixed set of rays (Fig. 6(a)), resulting in a set of polygons that correspond to piles in the input image, as shown in Fig. 6(b). Figure 6(c) presents the overlaid segmentation generated using SAM.

StarDist achieved a Hoover correct detection score of 0.494 and an over-segmentation score of 0.242. Both correctly detected, and over-segmented piles contributed to an overall accuracy of 0.736. These scores suggest that StarDist is more sensitive to noise and that the piles in the dataset are densely packed. Additionally, the internal structure of the piles can cause over-segmentation. For example, StarDist can split elongated cells into multiple star-convex subsets that do not span the entire pile²⁵. The StarDist network can make robust predictions of its distance field, but it fails to assemble accurate pile masks in some cases because some piles in the dataset are not well-approximated by star-convex polygons.

SAM achieved a remarkable Hoover correct detection score of 0.714 and an over-segmentation score of 0.226. The combined contribution of accurately detected and over-segmented piles resulted in an exceptional overall accuracy of 0.940. Zero-shot learning empowers a model to effectively process and respond to input data that it has not explicitly encountered during the training phase. This remarkable ability stems from the model's acquisition of a generalised understanding of the data rather than a mere collection of specific

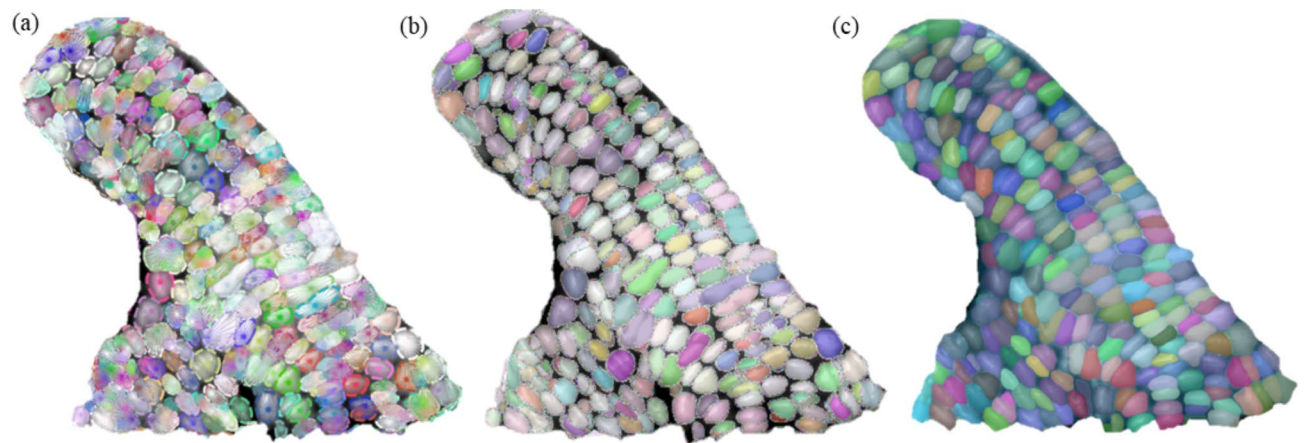


Fig. 6. (a) Star-convex polygons parameterised by the radial distances in hillshade representation of spoil pile. Segmented piles using (b) StarDist segmentation and (c) segment anything model (SAM).

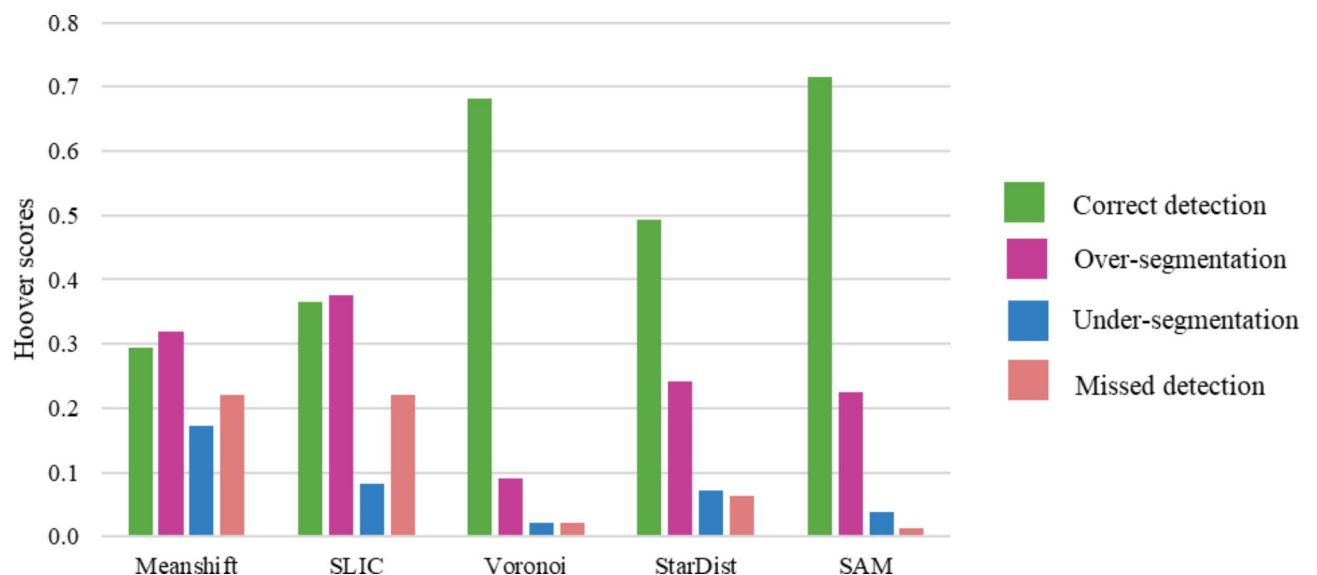


Fig. 7. Hoover metric scores for best performing traditional and deep learning-based segmentation approaches.

instances, culminating in accurate segmentation¹⁵. Zero-shot learning systems exhibit the remarkable ability to recognise piles they have never encountered before, relying on their understanding of underlying concepts and relationships.

Comparison of traditional and deep learning-based segmentation approaches

The performance of pile segmentation algorithms can be assessed through the Hoover score, a comprehensive metric that considers correct Detection, over-segmentation, under-segmentation, and missed detection (Fig. 7).

Among the algorithms evaluated, mean shift demonstrates a relatively balanced performance, with moderate scores across all metrics. SLIC exhibits a higher correct detection rate, indicating a better ability to accurately identify segments, but at the cost of increased over-segmentation. Voronoi, on the other hand, achieves an impressive correct detection score with minimal over-segmentation and under-segmentation, suggesting robust performance in delineating objects in the image. StarDist strikes a balance between correct detection and over-segmentation, showcasing a competitive overall performance. Notably, SAM stands out with the highest correct detection score, implying superior accuracy.

Upon closer examination, SAM appears to be particularly well-suited for spoil pile segmentation where precision is paramount. Its high correct detection score indicates a strong capability to identify segments accurately, while simultaneously minimising errors.

The selection of an image segmentation algorithm is contingent upon the specific requirements of the given task. Both SAM and Voronoi demonstrate promising outcomes in terms of overall performance. SAM exhibits notable strength, whereas Voronoi excels in achieving closer and more accurate detections.

Figures 8 and 9 provide a visual representation of segmentation performance. Visual assessment of segments from Figs. 8 and 9 align with the scores derived from Hoover metrics in Fig. 7.

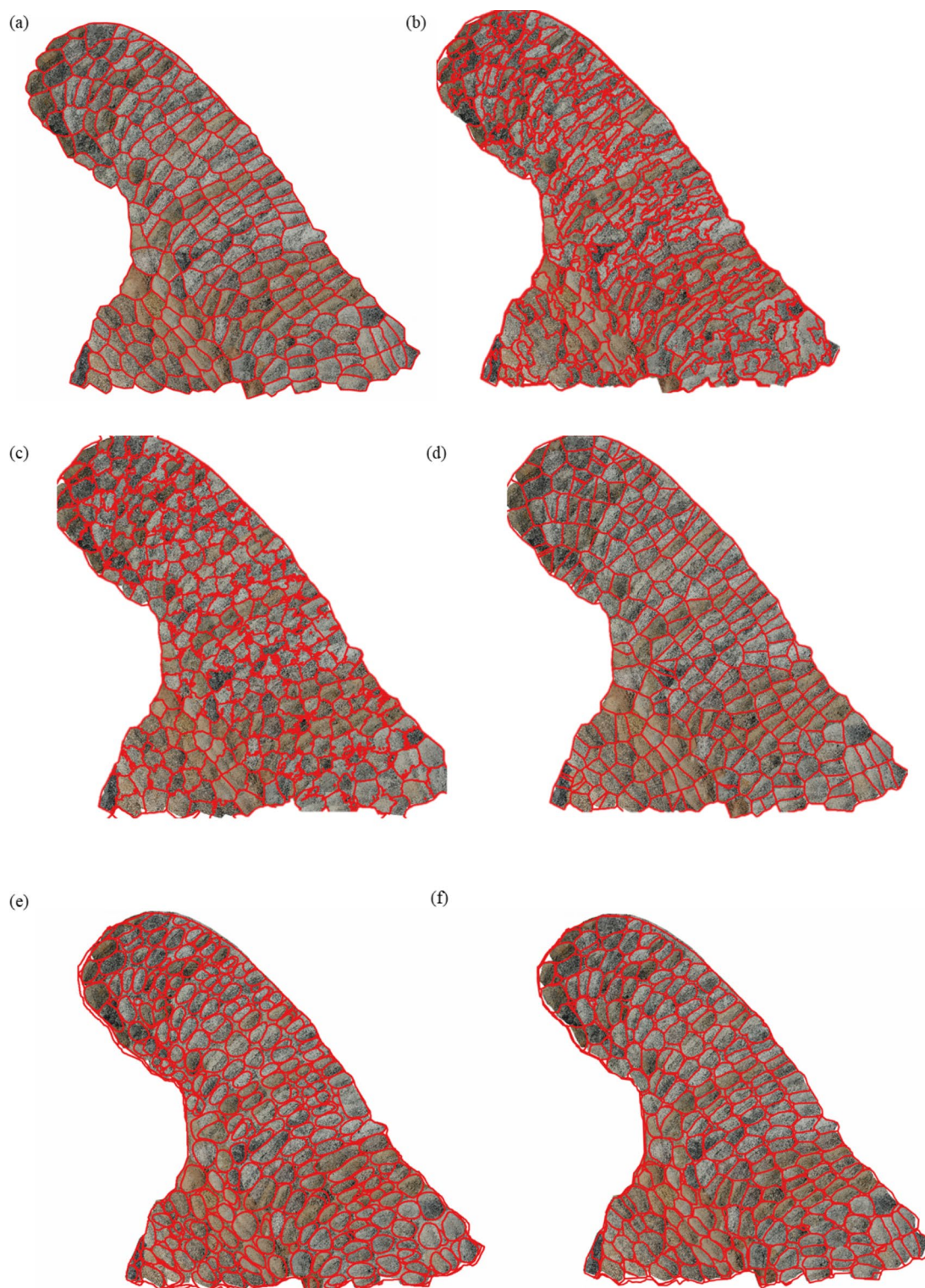


Fig. 8. (a) Reference segments and segmentation results of (b) mean shift algorithm, (c) simple linear iterative clustering, (d) Voronoi-based segmentation, (e) StarDist segmentation and (f) segment anything model.

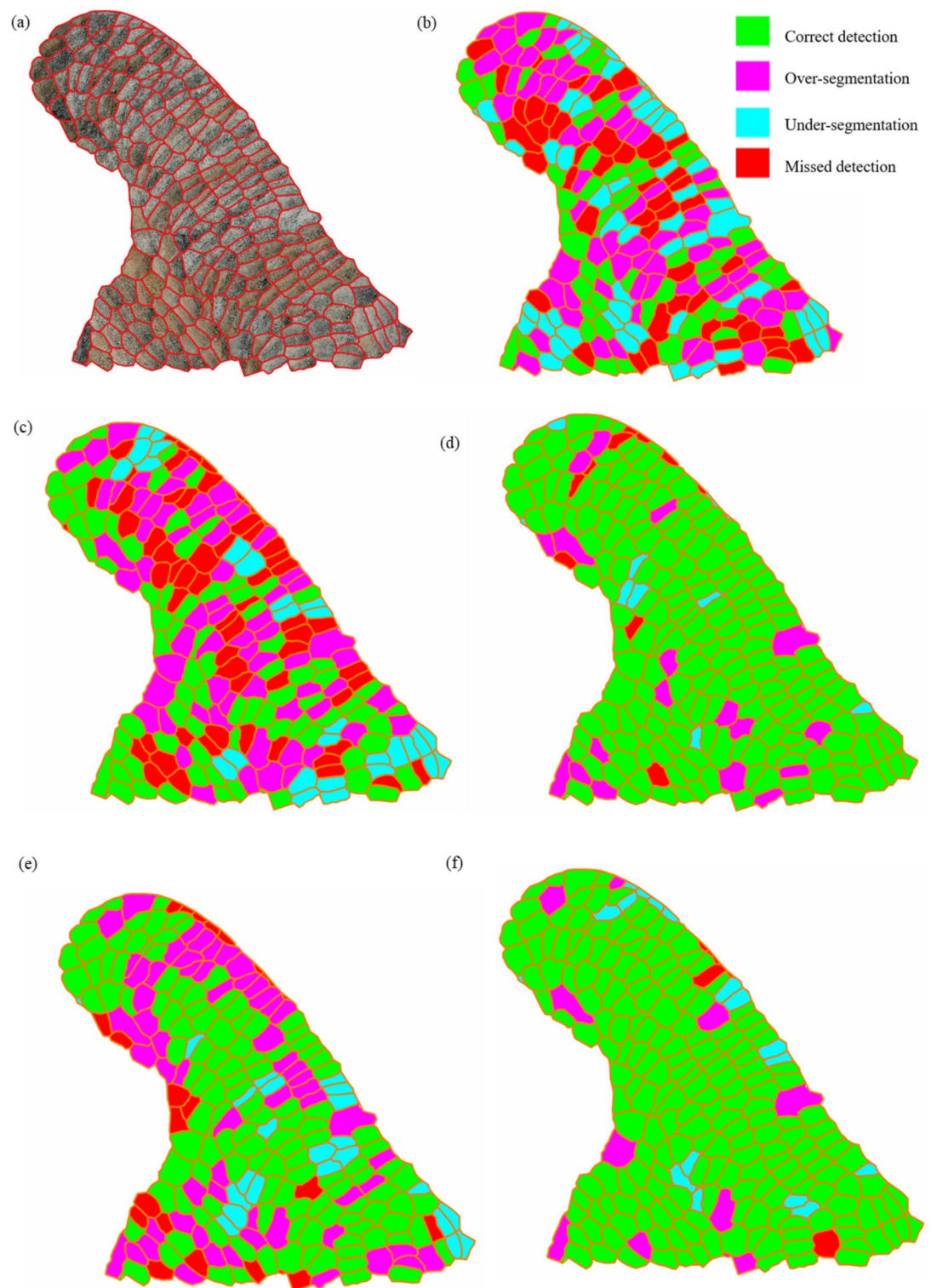


Fig. 9. (a) Reference segments and Hoover scores on reference for (b) mean shift algorithm, (c) simple linear iterative clustering and (d) Voronoi-based segmentation, (e) StarDist segmentation and (f) segment anything model

The assessment of segmentation methodologies indicates that morphology-based approaches (specifically Voronoi, StarDist, and SAM-based segmentation) outperform colour-based methods (mean shift and SLIC) in accurately delineating paddock tipped dumps. This superiority may stem from the susceptibility of colour-based approaches to external factors, such as shadows and variations in sun angles during image acquisition with UAVs. These factors can lead to the misclassification of pixels from the same class into distinct categories.

Notably, morphology-based segmentation, incorporating DSM and hillshade representations, remains less influenced by these external factors.

Among the evaluated segmentation approaches, Voronoi-based segmentation and SAM-based segmentation exhibit the highest correct detection rates at 0.681 and 0.714, respectively. It is noteworthy that the combination of correct detection and over-segmentation scores, where over-segmented polygons align with the geometry of reference segments and share their material properties, results in values of 0.771 and 0.940, respectively. This underscores the effectiveness of deep learning-based segmentation, with a particular emphasis on the performance of SAM in this context. This underscores the superior performance of deep learning approaches over traditional methods, showcasing their zero-shot learning capability. In this study, the primary objective was to use the pretrained capabilities of SAM and enhance its performance in spoil pile segmentation through fine-tuning. However, future work could involve a systematic analysis of the model's sensitivity to various training set and compositions to provide deeper insights into its adaptability and robustness.

Conclusions

The mining industry is undergoing a paradigm shift towards image-based analysis, driven by the imperative to alleviate the labour-intensive and inherently hazardous visual-based characterisation process. This study, centered on object-based image analysis, heralds a new era for the industry, paving the way for enhanced efficiency in spoil pile characterisation. However, the efficacy of advanced image processing techniques, such as object-based classification and feature extraction, hinges critically on segmentation quality. Consequently, identifying suitable spoil pile delineation algorithms becomes a pivotal step prior to classification. This study examines different ways to separate spoil piles, focusing on colour and morphology methods. The results indicate that using morphology-based segmentation outperforms colour-based methods for accurately defining spoil piles. Colour-based approaches exhibit a pronounced reliance on optical data, which is susceptible to perturbations induced by external factors such as sun angle and shadows in the undulating terrain of spoil piles, leading to misclassification of classes. In contrast, the morphology-based approach utilises DSM, which remains immune to these external illumination factors. Additionally, a comparison of traditional segmentation approaches with deep learning-based approaches underscores the effectiveness of deep learning-based approaches, particularly SAM, in the spoil pile delineation task. This efficacy stems from their zero-shot learning capability and inherent adaptability. Moreover, this investigation systematically explores the nuances associated with the selection of optimal parameters for segmentation methodologies applied to the delineation of spoil piles. The overarching objective is to augment the quality of segmentation outcomes. In the specific instance of SAM, four parameters were meticulously fine-tuned to achieve precision in spoil pile delineation, a process contingent upon the spatial resolution of the DSM. It is imperative to underscore that this fine-tuning approach holds transferability, as it can be extrapolated to other sites with paddock dumping practices. Moreover, adapting and fine-tuning these parameters for varying spatial resolutions is anticipated to contribute significantly to the enhancement of segmentation efficacy across different contexts. This study provides invaluable insights into segmentation approaches and corresponding parameters for spoil pile delineation, laying the foundation for automated image-based spoil characterisation in the impending future.

Data availability

The data that support the findings of this study are available from the mining site, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the corresponding author upon reasonable request and with permission of the mining site.

Received: 26 July 2024; Accepted: 23 October 2024

Published online: 25 March 2025

References

- Chugh, Y. P. & Behum, P. T. Coal waste management practices in the USA: an overview. *Int. J. Coal Sci. Technol.* **1**, 163–176. <https://doi.org/10.1007/s40789-014-0023-4> (2014).
- Shrivastava, S., Bhattacharjee, S. & Deb, D. Segmentation of mine overburden dump particles from images using Mask R CNN. *Sci. Rep.* **13**, doi: <https://doi.org/10.1038/s41598-023-28586-0> (2023).
- Banerjee, B., Raval, S. & Cullen, P. J. in 3rd International Future Mining Conference Sydney, Australia, (2015).
- Banerjee, B. P., Raval, S. & Cullen, P. UAV-hyperspectral imaging of spectrally complex environments. *Int. J. Remote Sens.* **41**, 4136–4159 (2020).
- Banerjee, B. P. & Raval, S. Mapping Sensitive Vegetation communities in Mining Eco-space using UAV-LiDAR. *Int. J. Coal Sci. Technol.* **9**, 1–16 (2022).
- Thiruchittampalam, S., Singh, S. K., Banerjee, B. P., Glenn, N. F. & Raval, S. Spoil characterisation using UAV-based optical remote sensing in coal mine dumps. *Int. J. Coal Sci. Technol.* **10**, 65. <https://doi.org/10.1007/s40789-023-00622-4> (2023).
- Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogrammetry Remote Sens.* **65**, 2–16 (2010).
- Dabov, P., Daud, M. & Olivotto, L. Revolutionizing urban mapping: deep learning and data fusion strategies for accurate building footprint segmentation. *Sci. Rep.* **14**, 13510. <https://doi.org/10.1038/s41598-024-64231-0> (2024).
- Michel, J., Grizonnet, M. & Canévet, O. in 2012 IEEE International Geoscience and Remote Sensing Symposium. 68–71.
- Thiruchittampalam, S., Banerjee, B. P., Singh, S. K., Glenn, N. F. & Raval, S. in IGARSS 2023–2023 IEEE International Geoscience and Remote Sensing Symposium. 2450–2453.
- De Luca, G. et al. Object-based land cover classification of Cork Oak Woodlands using UAV Imagery and Orfeo ToolBox. *Remote Sens.* **11**, 1238 (2019).
- Csillik, O. Fast segmentation and classification of very high Resolution Remote Sensing Data using SLIC superpixels. *Remote Sens.* **9** <https://doi.org/10.3390/rs9030243> (2017).
- Kaliman, S., Jayachandran, C., Rehfeldt, F. & Smith, A. S. Limits of Applicability of the Voronoi Tessellation determined by Centers of Cell Nuclei to epithelium morphology. *Front. Physiol.* **7** <https://doi.org/10.3389/fphys.2016.00551> (2016).

14. Stevens, M. et al. StarDist image segmentation improves circulating tumor cell detection. *Cancers (Basel)*. **14**, 2916 (2022).
15. Osco, L. P. et al. The segment anything model (SAM) for remote sensing applications: from zero to one shot. *Int. J. Appl. Earth Obs. Geoinf.* **124**, 103540. <https://doi.org/10.1016/j.jag.2023.103540> (2023).
16. Zheng, L., Zhang, J. & Wang, Q. Mean-shift-based color segmentation of images containing green vegetation. *Comput. Electron. Agric.* **65**, 93–98. <https://doi.org/10.1016/j.compag.2008.08.002> (2009).
17. Achanta, R. et al. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**, 2274–2282 (2012).
18. Alorf, A., K-Means, M., Shift & Clustering Algorithms, S. L. I. C. A *Comparison of Performance in Color-based Skin Segmentation* (University of Pittsburgh, 2017).
19. Yasuda, T. et al. Irradiation-injured brain tissues can self-renew in the absence of the pivotal tumor suppressor p53 in the medaka (*Oryzias latipes*) embryo. *J. Radiat. Res. (Tokyo)*. **57**, 9–15 (2016).
20. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. man. Cybernetics*. **9**, 62–66 (1979).
21. Schmidt, U., Weigert, M., Broaddus, C. & Myers, G. in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. (eds Alejandro F. Frangi et al.) 265–273 (Springer International Publishing).
22. Vazquez, K., Saraswathibhatla, A. & Notbohm, J. Effect of substrate stiffness on friction in collective cell migration. *Sci. Rep.* **12**, 2474. <https://doi.org/10.1038/s41598-022-06504-0> (2022).
23. Yu, Y., Sun, X. & Cheng, Q. Expert teacher based on foundation image segmentation model for object detection in aerial images. *Sci. Rep.* **13**, 21964. <https://doi.org/10.1038/s41598-023-49448-9> (2023).
24. Hoover, A. et al. An experimental comparison of range image segmentation algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.* **18**, 673–689 (1996).
25. Cutler, K. J. et al. Omnipose: a high-precision morphology-independent solution for bacterial cell segmentation. *Nat. Methods*. **19**, 1438–1448. <https://doi.org/10.1038/s41592-022-01639-4> (2022).

Acknowledgements

ACARP industry monitors, mine geotechnical engineers, mine geologists and mine surveyors are acknowledged for their assistance in collecting data for this project during the field campaigns.

Author contributions

Conceptualisation: S.T., B.P.B., N.F.G. and S.R.; Data curation: S.T.; Formal analysis: S.T.; Funding acquisition: S.R.; Investigation: S.T.; Methodology: S.T., B.P.B., N.F.G. and S.R.; Validation: : S.T., B.P.B., N.F.G. and S.R.; Visualisation: S.T.; Writing—original draft: S.T.; Supervision: B.P.B., N.F.G. and S.R.; review and editing: B.P.B., N.F.G. and S.R.

Funding

This work was supported by the Australian coal industry’s research program [grant number C29048].

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024