

# Enhancing Geological Mapping through Latent Diffusion Inpainting of Surface Infrastructure Artefacts in Remote Sensing Imagery

Mahsa Paknezhad <sup>1</sup>, Jack Maughan <sup>2</sup>, Tom Carmichael <sup>3</sup>

<sup>1</sup> [mahsapaknezhad@datarock.com.au](mailto:mahsapaknezhad@datarock.com.au), Senior Machine Learning Engineer, Datarock, Melbourne, VIC, Australia

<sup>2</sup> [jackmaughan@datarock.com.au](mailto:jackmaughan@datarock.com.au), Senior Data Scientist, Datarock, Melbourne, VIC, Australia

<sup>3</sup> [tomcarmichael@datarock.com.au](mailto:tomcarmichael@datarock.com.au), Chief Data Scientist, Datarock, Melbourne, VIC, Australia

Remote sensing data such as aerial and satellite imagery serve as valuable tools for mapping geological features. However, these datasets often present statistical limitations by capturing only the surface of the Earth. This limitation becomes pronounced when regions of interest, such as economic mineralization or specific geological formations are studied mainly because surface infrastructure artefacts, including mines, roads, dams, and drill pads, inadvertently influence the signature within the imagery. Such non-geological artefacts interfere with machine learning models when focused on geological responses, necessitating the masking or removal of these artefacts.

Recent advancements in machine learning and image processing demand comprehensive datasets devoid of missing data. To address this challenge, various imputation methods have been developed for inpainting missing spatial data. Among these methods, diffusion techniques emerge as modern, state-of-the-art approaches for inpainting data based on surrounding geological context.

In the field of computer vision and artificial intelligence (AI), existing inpainting models, including diffusion models, are predominantly designed for processing natural images composed of the standard red, green, and blue (RGB) colour channels. However, in remote sensing applications, image data often consists of more than just the three RGB bands. This expanded dataset may include numerous spectral bands, each providing valuable information about the geological features present.

Traditionally, one approach to address this complexity involves treating each spectral band as a separate greyscale image and training individual inpainting models for each band. While feasible, this method overlooks the potential benefits of leveraging the inter-band relationships inherent in multi-band imagery. A more effective strategy entails training a single inpainting model capable of processing all bands simultaneously. By doing so, the model can better exploit the intricate relationships between spectral bands to inpaint missing regions in a more realistic manner.

Despite the advantages of this unified approach, existing AI frameworks often lack support for images with more than three bands, limiting the applicability of traditional inpainting methods in geological contexts. To address this gap, we have developed an advanced inpainting model specifically tailored to handle images with an arbitrary number of spectral bands. Our model not only overcomes these technical limitations but also enhances the accuracy and versatility of inpainting tasks in geological imagery analysis.

This abstract underscores the significance of inpainting methodologies in enhancing the accuracy and reliability of geological mapping based on aerial and satellite imagery. By effectively removing surface infrastructure artefacts and inpainting missing spatial data, these techniques enable more precise prospectivity mapping and facilitate the advancement of geological understanding and exploration strategies in diverse landscapes.