PRINCIPAL COMPONENTS VERSUS AUTOENCODERS FOR DIMENSIONALITY REDUCTION: A CASE OF SUPER-RESOLVED OUTPUTS FROM PRISMA HYPERSPECTRAL MISSION DATA

K. Mishra, Benoit Vozel, R. D. Garg

PRINCIPAL COMPONENTS VERSUS AUTOENCODERS FOR DIMENSIONALITY REDUCTION: A CASE OF SUPER-RESOLVED OUTPUTS FROM PRISMA HYPERSPECTRAL MISSION DATA

K. Mishra [1*], B. Vozel [2], R. D. Garg [1]

[1] Geomatics Engineering Group, Civil Engineering Department, Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, 247667, India - (kmishra, rdgarg)@ce.iitr.ac.in
[2] MULTIP Research Group, Department IMAGE, Institut d'Electronique et des Technologies du numéRique (IETR) UMR CNRS 6164, Université de Rennes, F-22305 Lannion, France - benoit.vozel@univ-rennes.fr

KEY WORDS: PRISMA, Single-frame Super-resolution, Dimensionality Reduction, Artificial Intelligence, Quality Assessment

ABSTRACT:

This study attempts to solve these issues associated with hyperspectral (HS) data, i.e., coarse spatial resolution and high volume, by understanding the effect of deep learning and traditional dimensionality reduction on super-resolved products generated from the recently launched PRecursore IperSpettrale della Missione Applicativa (PRISMA) HS mission. Four single-frame super-resolution (SR) algorithms have been used to super-resolve a 30 m PRISMA scene of Ahmedabad, India and generate 15 m spatial resolution images with both spatial and spectral fidelity. Iterative back projection (IBP) and sparse representation (SIS) are the best and worst-performing SR algorithms following a comparative assessment and validation protocol. Next, denoising autoencoders and PCT computed using singular and eigenvalue decompositions have been executed on the original PRISMA, IBP and SIS-based super-resolved datasets. The resulting low-dimensional representations have been assessed to preserve the original dataset's topology using label-independent Lee and Verleysen's co-ranking matrix and loss of quality measure. Findings suggest that autoencoders are computationally expensive and require a higher neighbourhood size than PCT and its variants to produce a high-quality encoding. These insights remain significant for urban information extraction as there are few direct comparative assessments between machine learning-based linear and non-linear data compression methods in earlier studies.

1. INTRODUCTION

The advent of imaging spectroscopy in the form of airborne hyperspectral (HS) sensors in the 1980s and subsequently in spaceborne HS sensors at the beginning of the twenty-first century boosted the efforts in performing detailed mapping and monitoring of the Earth's resources and surface processes. Ameliorations in sensor optics and the need to ensure data continuity following Hyperion's decommissioning led to the launch of several HS spaceborne missions from 2018 onwards. Consequently, a huge amount of HS data is available nowadays with a wide swath, medium spatial resolution, and high signal-to-noise ratio (SNR), representing the five Vs of big data (Nguyen, 2019). Replete with similar information in multiple bands, feature redundancy hampers the data quality and hinders their usage in various thematic applications. Hence, there is a need to develop workflows focusing on retrieving spectral bands that contain the maximum information for the particular application, thereby highlighting the importance of dimensionality or spectral reduction techniques.

Such techniques help to avoid Hughes' phenomenon or the curse of dimensionality (Ma et al., 2013), due to which an exceptionally large number of points are needed for representation, and many machine learning algorithms underperform in retrieving precise content from HS data. Moreover, with a large number of dependent variables being recorded, there will be noise introduced due to erroneous measurements along with redundancy, leading to the increased need for storage, higher computation time, and substandard visualization in the absence of advanced tools (Xu et al., 2019). So, these techniques must be implemented as a pre-processing step to produce computationally efficient and accurate classification pipelines without information loss (S. H. Lee et al., 2018).

The most well-known technique is the principal component transform (PCT) (Jolliffe & Cadima, 2016), which, when used on HS data, generates components free of correlation and noise and possesses the maximum information. PCT-derived narrow-band indices from new HS sensors: DLR Earth Sensing Imaging Spectrometer (DESIS) (Krutz et al., 2019), PRecursore IperSpettrale della Missione Applicativa (PRISMA) (Pignatti et al., 2015) produce the most robust lithological maps in Rajasthan, India (Tripathi & Garg, 2021b). Tripathi & Garg (2021a) compare eigenvalue decomposition (EVD) and singular value decomposition (SVD) methods for PCT of Earth and lunar HS data. The compression rate and dimensionality reduction techniques are important while developing a classification workflow for HS images (Mantripragada et al., 2022). The last five years have seen the evolution of semisupervised and supervised algorithms in executing effective dimensionality reduction and image classification compared to conventional approaches like PCT (Harikiran & Reddy, 2022). PCT and its variants consider the image distribution linear, only possible in areas with homogeneity but not in locales with heterogeneity, such as an urban area (Mishra & Garg, 2023). Many non-linear techniques: locally linear embedding (Roweis & Saul, 2000), diffusion maps (Coifman & Lafon, 2006), sequential forward and backward selections (Wang et al., 2011), genetic algorithms (Chatterjee & Bhattacherjee, 2011) have been developed to solve these deficiencies.

*Corresponding author

Supervised deep learning strategies like convolutional neural networks (Hu et al., 2015) requiring many labelled samples to learn the complex patterns within the data have also been used for dimensionality reduction. However, each deep learning strategy takes care of certain properties of the original data and may be suitable for a particular task, not another. Tuning parameters and assumptions adopted by each strategy may lead the algorithm to produce diverse outputs each time corresponding to differing local optimal values of the objective function. As a result, the underlying structure of the input data may not be learnt properly and adequately represented by the inner layers of the neural network. Unsupervised non-linear neural network algorithms like autoencoders have shown superior performance in obtaining explainable structural mappings of the HS dataset in a compressed sub-space defined by the HS dataset's intrinsic dimensionality (Facco et al., 2017). Being independent of labels as in supervised learning, these representations can be deployed for further analyses or utilized on other scenes with varied heterogeneity (Sedhain et al., 2015). Priya et al. (2019) have found that autoencoders adequately capture the non-linear patterns within airborne HS data. However, unlike supervised learning, the output quality or the performance of two unsupervised methods cannot be compared with objective metrics like reconstruction error or classification accuracies.

Additionally, real-world spaceborne HS data have medium spatial resolution, which is insufficient for urban area mapping. Due to sensor optics design, the trade-off between detailed spatial and spectral information is difficult to address at the hardware stage. Resolution enhancement strategies such as super-resolution (SR), which attempts to retain the spectral information of a single or multiple low-resolution (LR) HS inputs while visualizing the scene at a higher spatial resolution, is a solution. Readers can refer to Aburaed et al. (2023) for more information on SR algorithm types and super-resolved products' quality assessment.

The present work aims to understand and analyze the impact of unsupervised linear and non-linear approaches in reducing the dimensionality of the newly launched PRISMA and super-resolved products generated from the original PRISMA dataset. Quality index based on the co-ranking matrix principle and loss of quality concept have been used for this purpose.

The remaining paper has been organized as follows. The region of investigation has been discussed in section 2. Section 3 describes the input dataset, software and machine configuration for executing the research. The detailed methodology is in section 4. Sections 5 and 6 discuss the research work's results and conclusions, respectively.

2. STUDY AREA

A 38.8 square kilometre (sq. km.) portion occupying the west and central zones of the Ahmedabad municipal corporation in Gujarat, India, has been chosen as the region of investigation. This region has a mixed land use containing residential, institutional and commercial areas in a planned manner. Informal settlements such as slums are also present, and every type of settlement density ranging from low to high is prevalent, too. As a result, the study area represents almost every type of building rooftop and pavement material used in the city along with natural features such as the bare ground and the dense green canopy in and around the Gujarat University, the Sabarmati River or the lakes in Vastrapur and Jivraj Park localities. Figure 1 below shows the study area on a false colour composite (FCC) of the PRISMA data.

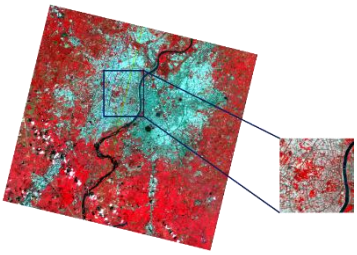

Figure 1. Study area on PRISMA FCC (Red (R): 37, Green (G): 24, Blue (B): 15)

3. DATASET, SOFTWARE AND TERMINAL SPECIFICATIONS

As specified in Section 1, data from the Italian Space Agency's (ASI's) newly launched HS mission has been used here. The sensor specifications are in Loizzo et al. (2018), while readers can refer to Mishra & Garg (2022) for the scene information.

The Environment for Visualising Images (ENVI) version 5.3 has been used for HS data pre-processing operations such as sensor error corrections, geometric correction and spatial or spectral subset. PCT using EVD of correlation or covariance matrix has also been executed in ENVI. The Matrix Laboratory (MATLAB) R2020b has been utilized for super-resolving the PRISMA data, estimating the dataset's intrinsic or virtual dimensionality, reducing its dimensionality by simple and deep autoencoders and PCT using singular value decomposition (SVD) of correlation or covariance matrix, and evaluating the quality of the dimensionally reduced outputs. R programming language has been used for stacking and assigning the coordinate projection system to the super-resolved output.

The research has been carried out on a single terminal with the central processing unit (CPU) specification: random access memory (RAM) = 64 GB, processor = Intel Xeon E5-1630 version 4, clock speed = 3.70 GHz.

4. METHODOLOGY

The methodology adopted for this research consists of HS data pre-processing, super-resolved output generation, performing dimensionality reduction and quality evaluation of dimensionally reduced outputs. These steps have been described in the following sub-sections.

4.1 PRISMA Data Pre-processing

The PRISMA dataset is a level-2D file, indicating that no atmospheric correction is required. The separate VNIR and SWIR data cubes are combined into a single file in standard ENVI format using the R programming language package 'prismaread' (Busetto & Ranghetti, 2020). Through a 1-degree polynomial affine transformation, the dataset is geometrically aligned to a terrain-corrected Hyperion dataset acquired in 2002 around the same time as the PRISMA data. The 'spatial subset' operation yields the study area in the dataset in square shape, having 209 samples and 209 lines. Manual inspection of each spectral band for haze, one or more than one column of no information or pixels having very low radiometric values led to the removal of 125 spectral bands out of 237 spectral bands in total.

4.2 Super-resolved Product Generation

The pre-processed PRISMA data is segregated into individual bands using the 'spectral subset' operation and fed into the SR algorithm as the LR input. Band-wise SR process execution has been performed to ascertain the algorithm's computation efficiency. Initially, four single-frame SR algorithms: iterative back projection (IBP) (Elad & Feuer, 1996; Irani & Peleg, 1991; Yang et al., 2014), sparse representations (SIS) (Zeyde et al., 2012), neighbourhood embedding – locally linear embedding (NELLE) (Chang et al., 2004), and gaussian process regression (GPR) (He & Siu, 2011) were applied 112 times for a scale factor (SF) = 2 to produce HR spectral bands at 15 m spatial resolution. These bands were stacked and given the coordinate projection system of the input dataset to generate the super-resolved output.

Readers can refer to the above-cited papers or Mishra & Garg (2022) for the algorithms' working and parameters adopted during implementation. The algorithms' performance was assessed using the comparative assessment and validation protocol proposed by Mishra & Garg (2022). IBP and SIS were the best and worst-performing SR algorithms among the selected algorithms. These super-resolved products and the original PRISMA dataset were chosen to assess the effect of dimensionality reduction.

4.3 Intrinsic Dimensionality Estimation

Before using PCT and autoencoders to reduce dimensionality, estimating the dataset's intrinsic dimensionality is important so that the target number of components needed to define the low dimensional feature space is known. The maximum likelihood estimation (MLE) (Levina & Bickel, 2004) algorithm has been utilized to obtain the intrinsic dimensionality of PRISMA and IBP and SIS-based super-resolved products. It is simple to execute and computationally efficient compared to existing approaches (Bruske & Sommer, 1998; Camastra & Vinciarelli, 2002; Costa & Hero, 2004; Roweis & Saul, 2000), offering the best bias and variance and assumes that every point's neighbourhood can be represented by a constant density linear feature space (Gomtsyan et al., 2019). However, this assumption may not hold for many real-world datasets, where each point may be represented by a non-linear manifold with a variable density, thereby prompting the development of feature-geometry-sensitive intrinsic dimensionality estimation algorithms.

Readers may refer to Gomtsyan et al. (2019) and Levina & Bickel (2004) for the mathematical derivation of the MLE descriptor. MLE's MATLAB implementation shown in 'MATLAB Toolbox for Dimensionality Reduction' (Van Der Maaten et al., 2009) has been followed here.

4.4 Dimensionality Reduction

PCT and autoencoders have been used to reduce the spectral dimensions of the original PRISMA dataset and the IBP and SIS-based super-resolved products derived from it. Here, EVD and SVD of covariance and correlation matrices derived from the input dataset have been used to determine the principal components (PC). Denoising single and deep autoencoders have also been utilized to visualize the input dataset in lower dimensions.

The following sub-sections describe these approaches.

4.4.1 PCT: This technique is named after Damodar Dharmananda Kosambi, Kari Karhunen and Michel Loève as the Kosambi-Karhunen-Loève transform (Sapatnekar, 2011). It points out similar and dissimilar patterns in multidimensional data using multivariate statistics. The focus is on obtaining the highest spectral variance by looking around for a new pair of orthogonal axes with their centre at the data's average value. The data is moved to a new projection system upon finding such a pair. EVD or SVD of the correlation or covariance matrix of the original image supports these steps. The difference between EVD and SVD lies in the type of vectors used to define the feature space. The former uses an eigenvector, which is a non-zero $n \times 1$ vector ϑ (Vadlamudi & Vadlamudi, 2018) such that,

$$\mathrm{P}\vartheta = \lambda\vartheta \quad (1)$$

where, $\mathrm{P} = n \times n$ square matrix
λ = eigenvalue

λ is a scalar and can be real or complex, and EVD holds for square matrices only. On the other hand, SVD uses right and left singular vectors μ and ω, respectively (Andrilli & Hecker, 2010) such that,

$$\mathrm{K} = \mathrm{M}\Delta\Omega^{\mathrm{T}} \quad (2)$$

where, $\mathrm{K} = m \times n$ rectangular matrix
$\mathrm{M} = m \times m$ orthogonal matrix consists of left singular vector set $\{\mu_1, \mu_2, \dots, \mu_m\}$ as its columns
$\Omega = n \times n$ orthogonal matrix consists of right singular vector set $\{\omega_1, \omega_2, \dots, \omega_n\}$ as its columns
$\Delta = m \times n$ diagonal matrix consists of all elements as 0, except elements at (i, i), which is a singular value σ_i such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k > \sigma_{k+1} = \cdots = \sigma_n = 0 \; \forall \; i \leq k$

SVD holds for rectangular or square matrices having real or complex elements. SVD's features include compressing imageries and retrieving optimum decomposition sub-rank estimates (Tripathi & Garg, 2021a). The resulting components from EVD or SVD are then ranked in descending order, and those with the highest values are selected to transform into a new feature space.

These components also do not correlate and are linearly mixed proportions of the original data's spectral bands in feature space with low dimensions. The highest amount of information is found in the first principal component, which has the highest variance (Jolliffe & Cadima, 2016). Even if the increasing PC number suggests lowering variance, it does not imply that a component with a smaller variance value will have less priority during regression (Jolliffe & Cadima, 2016). The noise in the data governs PCT's efficiency, and it is very much possible to have high-quality PCs at higher PC numbers (Hyvärinen & Oja, 2000). The reversal of both EVD and SVD makes it possible to obtain the original P and K, respectively.

As pointed out in Section 3, ENVI has been used to calculate the correlation and covariance matrices from the original PRISMA and super-resolved products and execute EVD on both correlation and covariance matrices. For SVD of both correlation and covariance matrices, MATLAB script[1] has been developed by the authors following the functions and syntax in the toolboxes: 'Open Source MATLAB Hyperspectral Toolbox' (Gerg, 2006) and 'MATLAB Toolbox for Dimensionality Reduction' (Van Der Maaten et al., 2009). The reduced dataset has been saved in ENVI standard format to facilitate easy visualization in ENVI.

[1]https://github.com/laserlight-cell/PCA-SVD

4.4.2 Autoencoders: According to Hinton & Salakhutdinov (2006), an unsupervised feed-forward neural network comprising an encoder $\mathcal{L}_E$ and a decoder $\mathcal{L}_D$ is called an autoencoder. $\mathcal{L}_E$ learns a low-dimensional hidden representation r_u of the input data A whereas $\mathcal{L}_D$ attempts to recreate the input from r_u such that $\widehat{\mathrm{A}} = \mathcal{L}_D(r_u)$ while simultaneously reducing the loss function

$$\mathcal{M}\left(\mathrm{A}, \mathcal{L}_D\left(\mathcal{L}_E(\mathrm{A})\right)\right) \quad (3)$$

where, $r_u = \mathcal{L}_E(\mathrm{A})$ and replacing $\mathcal{L}_D\left(\mathcal{L}_E(\mathrm{A})\right)$ in the above equation, by $\mathcal{L}_D(r_u)$ and subsequently by $\widehat{\mathrm{A}}$, loss function becomes $\mathcal{M}\left(\mathrm{A}, \widehat{\mathrm{A}}\right)$. This function measures the discrepancy between A and $\widehat{\mathrm{A}}$ and is generally a mean square error (Rasti et al., 2018). Let h_E and h_D be the weight matrices of $\mathcal{L}_E$ and $\mathcal{L}_D$, respectively, then the basic autoencoder's forward pass is

$$\widehat{\mathrm{A}} = \sigma_D\left(h_D\left(\sigma_E(h_E\mathrm{A})\right)\right) \quad (4)$$

where, σ_D and σ_E denote the output and latent layer's activation function, respectively. Readers must note that a basic or simple autoencoder consists of only one latent layer, which contains r_u and has d junctions such that $d < D$, where D is the number of junctions in the input and output layer. Although D should be equal in both layers preferably, asymmetry is also possible in the network architecture. Sigmoid activation functions are used in the input and output layers to understand the non-linear patterns in A, while the latent layer's activation function is linear. Since the loss function is unaffected by the characteristics of the latent layer, a regularization term must be introduced to warrant that all sorts of possible representations within A are learnt, and accurate reconstructions happen. Hence, the regularized loss function can be expressed as

$$\mathcal{M}_{regularized}\left(\mathrm{A}, \widehat{\mathrm{A}}\right) = \mathcal{M}\left(\mathrm{A}, \widehat{\mathrm{A}}\right) + \tau\mathcal{F}(r_u, h_E, h_D) \quad (5)$$

where, $\mathcal{F}(r_u, h_E, h_D)$ is a penalty function keeping track of the latent layer characteristics and weights of $\mathcal{L}_E$ and $\mathcal{L}_D$ and $\tau =$ tuning parameter controlling the regularization strength.

A simple autoencoder becomes a deep autoencoder when the latent layers increase in number and are usually odd, thereby allowing the autoencoder to address complex issues. However, the increment in the number of hidden layers leads to a reduction in the miscalculations sent back to the previous layer, implying regular weight updation in the latent layers adjoining the output layer. In contrast, the latent layers' weights near the input layer are seldom or never updated. Also known as the vanishing gradient problem (Roodschild et al., 2020), this issue prohibits learning very deep autoencoders and can be solved using a three-stage pre-training procedure (Hinton et al., 2006; Zheng & Zhao, 2020). Each layer from A to r_u is trained individually in the first stage using a restricted Boltzmann machine (Fischer & Igel, 2012) or a small denoising autoencoder (Vincent et al., 2008) as in the current research.

A small denoising autoencoder is nothing but a simple autoencoder that transforms A into a noisy version r_u in the latent layer during training via a stochastic function and then attempts to reconstruct the original data, i.e., obtain $\widehat{\mathrm{A}}$ from this noisy version using (3) and (4). The second stage is the inversing of the trained encoder layers to obtain the latent and output layers of the decoder. Supervised finetuning of the trained network by backpropagation forms the third phase. Spectral features generated from this pre-training perform better than the traditional information extraction strategies (Kong & Yan, 2019).

The primary drawback of autoencoders is that their training is time-consuming and highly dependent upon the number of data points in the dataset, the number of iterations and the number of weights. An increment in any of these factors results in increased training time as well as occupying increased memory space. Another point to remember when constructing an autoencoder is its discernment of the input dataset. Multiple varieties of autoencoders aiming to ameliorate the generalization and arrangement of source types have been introduced, such as variational (Liu et al., 2021), convolutional (Chen et al., 2017), denoising (Vincent et al., 2010), sparse (Shao et al., 2019) and adversarial autoencoders (Bao et al., 2021).

Again, the authors have developed a MATLAB script[2] for both simple and deep denoising autoencoders following the functions and syntax in the toolboxes: 'Open Source MATLAB Hyperspectral Toolbox' (Gerg, 2006) and 'MATLAB Toolbox for Dimensionality Reduction' (Van Der Maaten et al., 2009). The maximum number of iterations is 50, and noise level and τ have been assigned a default value of 0 for a fair comparison. As specified in sub-section 4.4.1, the dimensionally reduced representation has been stored in ENVI standard format for easy visualization in ENVI.

4.5 Quality Evaluation of Dimensionally Reduced Outputs

Visual examination of the reduced components or assessing their statistics to decide on the number of components to be retained in the dimensionally reduced output is a subjective way of establishing the efficiency of a spectral reduction technique. Moreover, if it is taken as the reference, a direct comparison with the original dataset does not make sense due to its high dimensional nature. Therefore, there is a need to develop measures to judge the low dimensional encoding's quality and assess the performance of the dimensionality reduction approaches.

Recently, rank measures based on the topology retention principle: agreement rate metric, trustworthiness and continuity and average relative rank errors have been suggested (Anowar et al., 2021; Gracia et al., 2014). This principle considers the geometric structure viewpoint, implying that neighbourhood points in a high-dimensional space should have their corresponding mappings in the low-dimensional space and the other way around. Introduced by Lee & Verleysen (2009), the quality metric grounded in the co-ranking matrix concept has been used here to evaluate the performance of low-dimensional representations generated by PCT and autoencoders. Several case studies across disciplines (Griparis et al., 2016; Silhan et al., 2019) deploy this metric for assessing dimensionality reduction by autoencoders. Ordinal distance changes are effectively encapsulated in a co-ranking matrix, substituting the column distances by their ranks. The rank comparison in the high and low dimensional expanses is done systematically. In an ideal scenario, non-zero elements are only available in diagonal places. The presence of a majority of non-zero elements in the upper triangle means that closely spaced points are far away. On the other hand, the lower triangle space filled with many non-zero elements indicates that faraway points are brought close to each other in the dimensionality reduction process.

Consider $\mathrm{A} = \{a_1, a_2, \ldots, a_n\} \in \mathbb{R}_\beta$ and $\widehat{\mathrm{A}} = \{\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_n\} \in \mathbb{R}_\alpha$ where $\mathbb{R}_\beta$ and $\mathbb{R}_\alpha$ are high and low-dimensional spaces such that $\alpha < \beta$. If Δ_{ij} is the distance between a_i and a_j in $\mathbb{R}_\beta$ and Λ_{ij}

[2]https://github.com/laserlight-cell/denoising-simple-and-deep-autoencoder-for-dimensionality-reduction

is the distance between $\hat{a}_i$ and $\hat{a}_j$ in $\mathbb{R}_\alpha$, then rank ς_{ij} of a_j concerning $a_i$in $\mathbb{R}_\beta$ is

$$\varsigma_{ij} = \left|\left\{\mathrm{K}|\Delta_{ik} < \Delta_{ij}\ or\ \left(\Delta_{ik} = \Delta_{ij}\ and\ 1 \le \mathrm{K} < \mathrm{j} \le \mathrm{N}\right)\right\}\right| \quad (6)$$

Likewise, rank υ_{ij} of $\hat{a}_j$ concerning $\hat{a}_i$ in $\mathbb{R}_\alpha$ is

$$\upsilon_{ij} = \left|\left\{\mathrm{K}|\Lambda_{ik} < \Lambda_{ij}\ or\ \left(\Lambda_{ik} = \Lambda_{ij}\ and\ 1 \le \mathrm{K} < \mathrm{j} \le \mathrm{N}\right)\right\}\right| \quad (7)$$

$$\therefore\ \Psi_{ij} = \upsilon_{ij} - \varsigma_{ij} \quad (8)$$

where, Ψ_{ij} is the rank miscalculation. The histogram of rank miscalculations is the co-ranking matrix Φ given by

$$\Phi_{kl} = \left|\left\{(i,j)|\varsigma_{ij} = \mathrm{K}\ and\ \upsilon_{ij} = 1\right\}\right| \quad (9)$$

Point pairs whose rank changes from $\mathbb{R}_\beta$ to $\mathbb{R}_\alpha$ are rank miscalculations and thus form non-zero elements in the upper or lower triangle of Φ. If $\varsigma_{ij} > \upsilon_{ij}$ for $\hat{a}_j$ then $\hat{a}_j$ is an intrusion, and vice-versa is called an extrusion. Φ can include several blocks corresponding to rank miscalculations, extrusions and intrusions. Mokbel et al. (2013) establish the Φ's unweighted sum as a quality metric Q_{NX}, which for P neighbourhood points is

$$Q_{NX}(P) = {}^{1}/_{PN} \sum_{k=1}^{P} \sum_{l=1}^{P} \Phi_{kl} \quad (10)$$

Subtracting the quality metric value from 1 can be utilized to understand the loss of the input data's geometrical structure during dimensionality reduction (Gracia et al., 2014). Computed over a pre-defined dimensional range, loss of quality is

$$Quality\ loss\ = 1 - Q_{NX}(P) \quad (11)$$

The domain of values falls between 0 and 1, with 0 denoting complete loss of geometrical structure and 1 meaning perfect retention. Hence, a smaller quality loss figure represents better preservation of geometry in the encoding.

The authors have developed a MATLAB script[3] for the quality evaluation of dimensionally reduced outputs following Lee & Verleysen (2009) and Mokbel et al. (2013).

5. RESULTS AND DISCUSSION

MLE gave the intrinsic dimensionality as 11, implying that the low-dimensional feature space will have 11 dimensions. Therefore, keeping the spatial dimension identical, the encoding in 11 spectral dimensions is compared with the original 112 spectral dimensions using Q_{NX}. It is noteworthy that even if the spatial resolution increases, the computation time for determining Q_{NX} remains almost constant, ranging between 13.6 to 14.7 seconds, depending upon the dataset. This is because the spatial dimension remains constant, consisting of 209 scan lines and 209 samples, irrespective of the increase in the spatial resolution. In the case of 15 m spatial resolution datasets, the central portion consisting of all possible distinct information classes in the scene has been spatially subsetted for computing Q_{NX} and $Quality\ loss$. The entire quality evaluation process takes about 1000 to 1100 seconds on average, irrespective of the dimensionality reduction algorithm, with the determination of the co-ranking matrix (about 693 – 705 seconds) taking double the amount of time than the determination of the ranking matrix (346 – 353 seconds).

Figure 2 a-c shows Q_{NX} for different values of K neighbouring points for PCT and autoencoder-based dimensionality reduction of PRISMA and its super-resolved outputs, respectively. Figure 2 d-f shows $Quality\ loss$ in a similar manner. For PRISMA data, the quality measure is higher at higher K values, rising to above 0.9 at values > 40000 for SVD-based PCT and autoencoder-based reductions. Their values remain almost identical. In contrast, EVD-based PCT values are above 0.95 for $K = 500$ and rise to above 0.99 with an increase in K, implying that SVD-based PCT of covariance or correlation matrix and deep and simple autoencoders give a better dimensionality reduction than EVD-based PCT for PRISMA. IBP and SIS-based super-resolved products show similar patterns as the original PRISMA data for PCT and autoencoders. $Quality\ loss$ figures also echo similar trends as Q_{NX}, indicating topology preservation at lower values of K for each method for PRISMA and the super-resolved outputs.

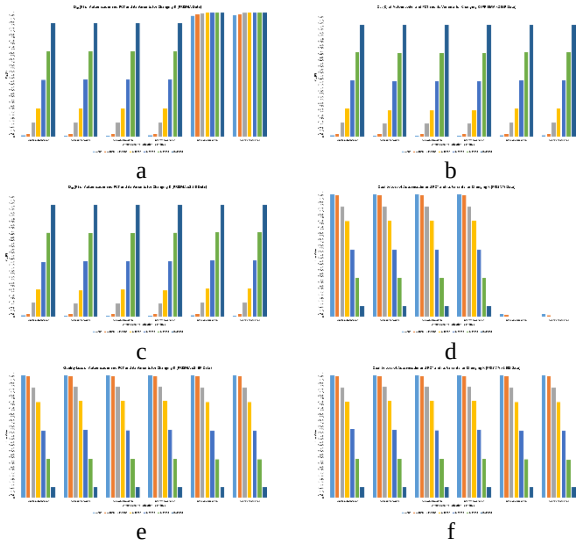

Figure 2. $\mathbf{Q_{NX}}$ and **Quality loss**, respectively, for PRISMA (a and d), IBP-based super-resolved product (b and e) and SIS-based super-resolved product (c and f) for K = 500 to 40000.

With little difference between PCT and autoencoder representations, the only distinguishing factor is the computation time. Both SVD and EVD-based PCT happen within a fraction of a second for the original PRISMA data. For super-resolved outputs, the amount of information increases the computation time to about 1.5 seconds. On the contrary, simple and deep autoencoders are computationally expensive, with the simple autoencoder taking about 127 seconds to learn the latent representation in a single run of 50 iterations and reducing the mean squared error from 0.45 to 0.0012 on average and the deep autoencoder taking roughly 2000 seconds for learning the latent representation in three runs of 50 iterations and reducing the mean squared error within identical limits. These times increase proportionately with the increase in scale factor corresponding to above 400 and 5000 seconds, although the network size remains constant, with the input layer having 140 nodes and the output layer having 11 nodes.

6. CONCLUSION

The current research evaluates SVD and EVD-based PCT and simple and deep autoencoder-based dimensionality reduction of PRISMA and its super-resolved products. Comparisons based on label-independent co-ranking matrix-based quality metric, quality loss concept and computation time reveal almost identical patterns for all approaches in the super-resolved outputs. Excellent topology preservation is apparent at K< 500, except in EVD-based PCT for the original PRISMA data. However, both the quality evaluation and autoencoder-based dimensionality reduction remain computationally expensive, raising doubts

[3]https://github.com/laserlight-cell/quality-evaluation-of-dimensionally-reduced-output

about the applicability of such processes when dealing with huge amounts of data and limited computation power. Moreover, the scale factor does not have any role to play in the autoencoder network size or the latent representation learnt from the input, indicating that super-resolved products can be treated similarly to the original HS data, albeit at a higher spatial scale.

Future work will look at the applicability of kernel-PCT, independent component analyses and MNF of the super-resolved outputs at higher scale factors and then proceed to automated endmember extraction.

ACKNOWLEDGEMENTS

The authors would like to express their heartfelt gratitude to ASI for making the PRISMA dataset freely available to the scientific community for research.

REFERENCES

Aburaed, N., Alkhatib, M. Q., Marshall, S., Zabalza, J., Al Ahmad, H., 2023: A Review of Spatial Enhancement of Hyperspectral Remote Sensing Imaging Techniques. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *16*, 2275–2300. https://doi.org/10.1109/JSTARS.2023.3242048

Andrilli, S., Hecker, D., 2010: Vectors and Matrices. In S. Andrilli & D. Hecker (Eds.), *Elementary Linear Algebra* (pp. 1–77). Academic Press. https://doi.org/10.1016/b978-0-12-374751-8.00001-9

Anowar, F., Sadaoui, S., Selim, B., 2021: Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE). *Computer Science Review*, *40*, 100378. https://doi.org/10.1016/j.cosrev.2021.100378

Bao, Z., Xue, R., Jin, Y., 2021: Image scrambling adversarial autoencoder based on the asymmetric encryption. *Multimedia Tools and Applications*, *80*(18), 28265–28301. https://doi.org/10.1007/s11042-021-11043-3

Bruske, J., Sommer, G., 1998: Intrinsic dimensionality estimation with optimally topology preserving maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(5), 572–575. https://doi.org/10.1109/34.682189

Busetto, L., Ranghetti, L., 2020: *prismaread: A tool for facilitating access and analysis of PRISMA L1/L2 hyperspectral imagery v1.0.0*. https://doi.org/https://doi.org/10.5281/zenodo.4019081

Camastra, F., Vinciarelli, A., 2002: Estimating the intrinsic dimension of data with a fractal-based method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(10),1404–1407.

Chang, H., Yeung, D. Y., Xiong, Y., 2004: Super-resolution through neighbor embedding. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *1*, I–I. https://doi.org/10.1109/cvpr.2004.1315043

Chatterjee, S., Bhattacherjee, A., 2011: Genetic algorithms for feature selection of image analysis-based quality monitoring model: An application to an iron mine. *Engineering Applications of Artificial Intelligence*, *24*(5), 786–795. https://doi.org/10.1016/j.engappai.2010.11.009

Chen, M., Shi, X., Zhang, Y., Wu, D., Guizani, M., 2017: Deep Feature Learning for Medical Image Analysis with Convolutional Autoencoder Neural Network. *IEEE Transactions on Big Data*, *7*(4), 750–758. https://doi.org/10.1109/tbdata.2017.2717439

Coifman, R. R., Lafon, S., 2006: Diffusion maps. *Applied and Computational Harmonic Analysis*, *21*(1), 5–30. https://doi.org/10.1016/j.acha.2006.04.006

Costa, J. A., Hero, A. O., 2004: Geodesic entropic graphs for dimension and entropy estimation in manifold learning. *IEEE Transactions on Signal Processing*, *52*(8), 2210–2221. https://doi.org/10.1109/TSP.2004.831130

Elad, M., Feuer, A., 1996: Super-resolution reconstruction of an image. *Proceedings of 19th Convention of Electrical and Electronics Engineers in Israel*, 391–394. https://doi.org/10.1109/EEIS.1996.566997

Facco, E., d'Errico, M., Rodriguez, A., Laio, A., 2017: Estimating the intrinsic dimension of datasets by a minimal neighborhood information. *Scientific Reports*, *7*(1), 12140. https://doi.org/10.1038/s41598-017-11873-y

Fischer, A., Igel, C., 2012: An introduction to restricted Boltzmann machines. In L. Alvarez, M. Mejail, L. Gomez, & J. Jacobo (Eds.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 7441 LNCS* (pp. 14–36). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-33275-3_2

Gerg, I., 2006: *Open Source MATLAB Hyperspectral Toolbox*. https://github.com/isaacgerg/matlabHyperspectralToolbox

Gomtsyan, M., Mokrov, N., Panov, M., Yanovich, Y., 2019: Geometry-Aware Maximum Likelihood Estimation of Intrinsic Dimension. *Proceedings of Machine Learning Research*, *101*, 1126–1141.

Gracia, A., González, S., Robles, V., Menasalvas, E., 2014: A methodology to compare Dimensionality Reduction algorithms in terms of loss of quality. *Information Sciences*, *270*, 1–27. https://doi.org/10.1016/j.ins.2014.02.068

Griparis, A., Faur, D., Datcu, M., 2016: A dimensionality reduction approach to support visual data mining: Co-ranking-based evaluation. *IEEE International Conference on Communications*, *2016-August*, 391–394. https://doi.org/10.1109/ICComm.2016.7528271

Harikiran, J., Reddy, T., 2022: An outlook: machine learning in hyperspectral image classification and dimensionality reduction techniques. *Journal of Spectral Imaging*, *11*(1), a1. https://doi.org/DOI:10.1255/jsi.2022.a1

He, H., Siu, W.-C., 2011: Single image super-resolution using Gaussian process regression. *CVPR 2011*, 449–456. https://doi.org/10.1109/CVPR.2011.5995713

Hinton, G. E., Osindero, S., Teh, Y.-W., 2006: A Fast Learning Algorithm for Deep Belief Nets. *Neural Comput.*, *18*(7), 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527

Hinton, G. E., Salakhutdinov, R. R., 2006: Reducing the Dimensionality of Data with Neural Networks. *Science*, *313*(5786), 504–507. https://doi.org/10.1126/science.1127647

Hu, W., Huang, Y., Wei, L., Zhang, F., Li, H., 2015: Deep Convolutional Neural Networks for Hyperspectral Image Classification. *Journal of Sensors*, *2015*, 258619. https://doi.org/10.1155/2015/258619

Hyvärinen, A., Oja, E. 2000: Independent component analysis: algorithms and applications. *Neural Networks*, *13*(4), 411–430. https://doi.org/https://doi.org/10.1016/S0893-6080(00)00026-5

Irani, M., Peleg, S., 1991: Improving resolution by image registration. *CVGIP: Graphical Models and Image Processing*, *53*(3), 231–239. https://doi.org/10.1016/1049-9652(91)90045-L

Jolliffe, I., Cadima, J., 2016: Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*,*374*(2065),20150202.https://doi.org/10.1098/rsta.2015.0202

Kong, D., Yan, X., 2019: Novel Model Based on Stacked Autoencoders with Sample-Wise Strategy for Fault Diagnosis. *Mathematical Problems in Engineering*, *2019*, 8985657. https://doi.org/10.1155/2019/8985657

Krutz, D., Müller, R., Knodt, U., Günther, B., Walter, I., Sebastian, I., Säuberlich, T., Reulke, R., Carmona, E., Eckardt, A., Venus, H., Fischer, C., Zender, B., Arloth, S., Lieder, M., Neidhardt, M., Grote, U., Schrandt, F., Gelmi, S., Wojtkowiak, A., 2019: The instrument design of the DLR earth sensing imaging spectrometer (DESIS). In *Sensors (Switzerland)* (Vol. 19, Issue 7). https://doi.org/10.3390/s19071622

Lee, J. A., Verleysen, M., 2009: Quality assessment of dimensionality reduction: Rank-based criteria. *Neurocomputing*, *72*(7),1431–1443. https://doi.org/https://doi.org/10.1016/j.neucom.2008.12.017

Lee, S. H., Kim, D. H., Song, B. C., 2018: Self-supervised knowledge distillation using singular value decomposition. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11210 LNCS*, 339–354. https://doi.org/10.1007/978-3-030-01231-1_21

Levina, E., Bickel, P., 2004: Maximum Likelihood Estimation of Intrinsic Dimension. In L. Saul, Y. Weiss, & L. Bottou (Eds.), *Advances in Neural Information Processing Systems* (Vol. 17). MIT Press. https://proceedings.neurips.cc/paper_files/paper/2004/file/74934548253bcab8490ebd74afed7031-Paper.pdf

Liu, Z. S., Siu, W. C., Wang, L. W., 2021: Variational autoencoder for reference based image super-resolution. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 516–525. https://doi.org/10.1109/CVPRW53098.2021.00063

Loizzo, R., Guarini, R., Longo, F., Scopa, T., Formaro, R., Facchinetti, C., Varacalli, G., 2018: Prisma: The Italian Hyperspectral Mission. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 175–178. https://doi.org/10.1109/IGARSS.2018.8518512

Ma, W., Gong, C., Hu, Y., Meng, P., Xu, F., 2013: The Hughes phenomenon in hyperspectral classification based on the ground spectrum of grasslands in the region around Qinghai Lake. *Proc.SPIE*, *8910*, 89101G. https://doi.org/10.1117/12.2034457

Mantripragada, K., Dao, P. D., He, Y., Qureshi, F. Z., 2022: The effects of spectral dimensionality reduction on hyperspectral pixel classification: A case study. *PLOS ONE*, *17*(7), e0269174. https://doi.org/10.1371/journal.pone.0269174

Mishra, K., Garg, R. D., 2022: Exploring single-frame super-resolution on real-world Hyperion and PRISMA datasets of an urban area in a developing nation. *International Journal of Remote Sensing*, *43*(12), 4569–4607. https://doi.org/10.1080/01431161.2022.2114109

Mishra, K., Garg, R. D., 2023: Evaluating Dimensionality Reduction Approaches on Erstwhile Hyperion and Newly Launched PRISMA Datasets. In M. Ben Ahmed, A. A. Boudhir, D. Santos, R. Dionisio, & N. Benaya (Eds.), *Innovations in Smart Cities Applications Volume 6* (pp. 36–45). Springer International Publishing.

Mokbel, B., Lueks, W., Gisbrecht, A., Hammer, B., 2013: Visualizing the quality of dimensionality reduction. *Neurocomputing*,*112*,109–123. https://doi.org/10.1016/j.neucom.2012.11.046

Nguyen, T. L., 2019: A Framework for Five Big V's of Big Data and Organizational Culture in Firms. *Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018*, 5411–5413. https://doi.org/10.1109/BigData.2018.8622377

Pignatti, S., Acito, N., Amato, U., Casa, R., Castaldi, F., Coluzzi, R., De Bonis, R., Diani, M., Imbrenda, V., Laneve, G., Matteoli, S., Palombo, A., Pascucci, S., Santini, F., Simoniello, T., Ananasso, C., Corsini, G., Cuomo, V., 2015: Environmental products overview of the Italian hyperspectral prisma mission: The SAP4PRISMA project. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *2015-Novem*, 3997–4000. https://doi.org/10.1109/IGARSS.2015.7326701

Priya, S., Ghosh, R., Bhattacharya, B. K., 2019: Non-linear autoencoder based algorithm for dimensionality reduction of airborne hyperspectral data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(3/W6), 593–598. https://doi.org/10.5194/isprs-archives-XLII-3-W6-593-2019

Rasti, B., Scheunders, P., Ghamisi, P., Licciardi, G., Chanussot, J., 2018: Noise Reduction in Hyperspectral Imagery: Overview and Application. In *Remote Sensing* (Vol. 10, Issue 3). https://doi.org/10.3390/rs10030482

Roodschild, M., Gotay Sardiñas, J., Will, A., 2020: A new approach for the vanishing gradient problem on sigmoid activation. *Progress in Artificial Intelligence*, *9*(4), 351–360. https://doi.org/10.1007/s13748-020-00218-y

Roweis, S. T., Saul, L. K., 2000: Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science*, *290*(5500), 2323–2326. https://doi.org/10.1126/science.290.5500.2323

Sapatnekar, S. S., 2011: Overcoming Variations in Nanometer-Scale Technologies. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, *1*(1), 5–18. https://doi.org/10.1109/JETCAS.2011.2138250

Sedhain, S., Menon, A. K., Sanner, S., Xie, L., 2015: AutoRec: Autoencoders Meet Collaborative Filtering. *Proceedings of the 24th International Conference on World Wide Web*, 111–112. https://doi.org/10.1145/2740908.2742726

Shao, Z., Wang, L., Wang, Z., Deng, J., 2019: Remote Sensing Image Super-Resolution Using Sparse Representation and Coupled Sparse Autoencoder. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(8), 2663–2674. https://doi.org/10.1109/JSTARS.2019.2925456

Silhan, T., Oehmcke, S., Kramer, O., 2019: Evolution of Stacked Autoencoders. *2019 IEEE Congress on Evolutionary Computation, CEC 2019 - Proceedings*, 823–830. https://doi.org/10.1109/CEC.2019.8790182

Tripathi, P., Garg, R. D., 2021a: Comparative Analysis of Singular Value Decomposition and Eigen Value Decomposition Based Principal Component Analysis for Earth and Lunar Hyperspectral Image. *Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote Sensing*, *2021-March*, 1–5. https://doi.org/10.1109/WHISPERS52202.2021.9483978

Tripathi, P., Garg, R. D., 2021b: Feature extraction of DESIS and PRISMA hyperspectral remote sensing datasets for geological applications. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLIV-M-3–2*, 169–173.https://doi.org/10.5194/isprs-archives-XLIV-M-3-2021-169-2021

Vadlamudi, C. V., Vadlamudi, S. P. D., 2018: Mathematical Essentials. In V. N. Gudivada & C. R. B. T.-H. of S. Rao (Eds.), *Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications* (Vol. 38, pp. 53–73).Elsevier. https://doi.org/10.1016/bs.host.2018.07.008

Van Der Maaten, L. J. P., Postma, E. O., Van Den Herik, H. J., 2009: Dimensionality Reduction: A Comparative Review. *Journal of Machine Learning Research*, *10*, 1–41. https://doi.org/10.1080/13506280444000102

Vincent, P., Larochelle, H., Bengio, Y., Manzagol, P.-A., 2008: Extracting and Composing Robust Features with Denoising Autoencoders. *Proceedings of the 25th International Conference on Machine Learning*, 1096–1103. https://doi.org/10.1145/1390156.1390294

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P. A., 2010: Stacked denoising autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. *Journal of Machine Learning Research*, *11*, 3371–3408.

Wang, L., Shen, C., Hartley, R., 2011: On the Optimality of Sequential Forward Feature Selection Using Class Separability Measure. *2011 International Conference on Digital Image Computing: Techniques and Applications*, 203–208. https://doi.org/10.1109/DICTA.2011.41

Xu, H., Zhang, H., He, W., Zhang, L., 2019: Superpixel-based spatial-spectral dimension reduction for hyperspectral imagery classification. *Neurocomputing*, *360*, 138–150. https://doi.org/10.1016/j.neucom.2019.06.023

Yang, C.-Y., Ma, C., Yang, M.-H., 2014: Single-Image Super-Resolution: A Benchmark. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science* (pp. 372–386). Springer International Publishing.

Zeyde, R., Elad, M., Protter, M., 2012: On single image scale-up using sparse-representations. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6920 LNCS*, 711–730. https://doi.org/10.1007/978-3-642-27413-8_47

Zheng, S., Zhao, J., 2020: A new unsupervised data mining method based on the stacked autoencoder for chemical process fault diagnosis. *Computers & Chemical Engineering*, *135*, 106755.https://doi.org/10.1016/j.compchemeng.2020.106755