



# Hybrid Adaptive Prediction Mechanisms with Multilayer Propagation Neural Network for Hyperspectral Image Compression

Rui Xiao<sup>(✉)</sup> and Manoranjan Paul

School of Computing and Mathematics, Charles Sturt University,  
Bathurst, Australia

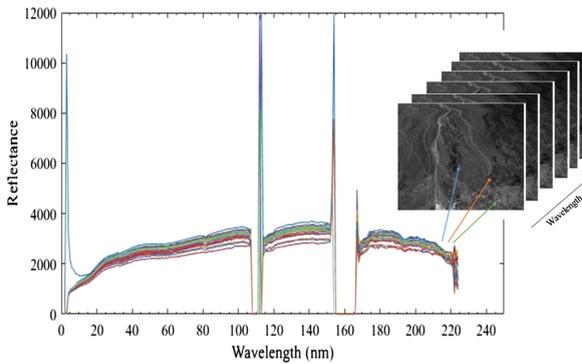
{rxiao, mpaul}@csu.edu.au

**Abstract.** Hyperspectral (HS) image is a three dimensional data image where the 3<sup>rd</sup> dimension carries the wealth of spectrum information. HS image compression is one of the areas that has attracted increasing attention for big data processing and analysis. HS data has its own distinguishing feature which differs with video because without motion, also different with a still image because of redundancy along the wavelength axis. The prediction based method is playing an important role in the compression and research area. Reflectance distribution of HS based on our analysis indicates that there is some nonlinear relationship in intra-band. The Multilayer Propagation Neural Networks (MLPNN) with backpropagation training are particularly well suited for addressing the approximation function. In this paper, an MLPNN based predictive image compression method is presented. We propose a hybrid Adaptive Prediction Mechanism (APM) with MLPNN model (APM-MLPNN). MLPNN is trained to predict the succeeding bands by using current band information. The purpose is to explore whether MLPNN can provide better image compression results in HS images. Besides, it uses less computation cost than a deep learning model so we can easily validate the model. We encoded the weights vector and the bias vector of MLPNN as well as the residuals. That is the only few bytes it then sends to the decoder side. The decoder will reconstruct a band by using the same structure of the network. We call it an MLPNN decoder. The MLPNN decoder does not need to be trained as the weights and biases have already been transmitted. We can easily reconstruct the succeeding bands by the MLPNN decoder. APM constrained the correction offset between the succeeding band and the current spectral band in order to prevent HS image being affected by large predictive biases. The performance of the proposed algorithm is verified by several HS images from Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) reflectance dataset. MLPNN simulation results can improve prediction accuracy; reduce residual of intra-band with high compression ratio and relatively lower bitrates.

**Keywords:** Hyperspectral image · MLP neural network · BP neural network  
Image coding · Data compression · Remote sensing · Inter-band prediction

## 1 Introduction

HS image combines the power of digital imaging and spectroscopy. By using push broom-scanning mode, HS camera sensor scans objects line by line simultaneously in a narrow spectral band. Hence every pixel in the HS image contains a wide range of spectrum and can be used to characterize the objects in the scene with great precision and detail. HS images are beyond human vision ability, can provide moisture content, texture, reflectance and other external quality characteristics of diverse samples. This comes at a price that HS image has a big data set and high redundancy. The extensive implementation of HS imaging is well founded in both the civilian and military field, such as satellite/airborne based remote sensing (NASA's Airborne Visible/Infrared Imaging Spectrometer 2017), target detection (Cheng and Han 2016; Makki et al. 2017), safety and quality inspection, classification as well as quality control in food and agriculture, (Cheng et al. 2017; Park and Lu 2015) and lab applications etc. In Fig. 1, an example of an HS image, namely cuprite from NASA's AVIRIS dataset is shown. This HS image carries 224 contiguous channels with wavelengths from 0.4–2.5  $\mu\text{m}$ . It shows that each pixel  $P = (p_1, p_2, \dots, p_n)$  is represented as a vector with 224 elements. Because carrying wealthy spectral information, the HS image data becomes huge. Compression of HS image has gained increasing interest.



**Fig. 1.** An HS image, cuprite demonstrates reflectance values curve from randomly picked pixels.

HS image compression methods are different with traditional compression methods as spectral redundancy needs to be removed; it also distinguishes itself from videos as HS data has no motion. HS image has its unique structure, each pixel along the spectral axis is characterised by nonlinearity, continuity and similarity. Especially similarity trend in neighbouring bands of the spectrum can be extracted as learning samples to train a neural network.

Compression techniques can broadly be grouped into two main categories: lossless and lossy compression methods, depending on whether the original image can be precisely re-generated from the compression data (Motta et al. 2006). Or according to different compression technical characteristics, there is transform-based compression,

prediction-based compression or vector quantization-based compression. Dictionary-based schemes such as Lempel-Ziv-Welch (LZW) compression algorithm are also prominent (King et al. 2014). Statistical-based schemes require distribution knowledge where the compression takes place based on the frequency of input characters. All these methods are based on lossless compression methods. The most well-known statistical-based algorithms are Huffman Coding (King et al. 2014) and Arithmetic Coding (Howard and Vitter 1992; Sasilal and Govindan 2013). Moreover, another type compression technique is based on LookUp Tables (LUT) (Aiazzi et al. 2009; Mielikainen and Toivanen 2008), which is also lossless type compression method. The LUT searches the previous band for a pixel equal to the current band in the same position called a predictor. The predictor is used as a key to search LUT to speed up the search process.

Transform-based technique, such as the Pairwise Orthogonal Transform (POT), also called multiple pairwise PCA (Amrani et al. 2016) is one of the spectral transforms that an image is transformed using multiple pairwise operations instead of a single transform. It overcomes the problem of KLT, such as bit depth expansion, lack of scalability and reduced memory requirements. (Shahriyar et al. 2016).

Prediction-based technique is an important research direction for compression coding. This type of method often uses a mathematical model to predict pixel values and encode only their prediction residuals (Conoscenti et al. 2016).

Zhao et al. (2016) and Zhu et al. (2015) introduce an algorithm based on intra-band prediction and inter-band fractal encoding. In this method, HS bands are partitioned into several Groups Of Bands (GOBs). The authors apply intra-band prediction to the first band then apply in each GOB. The hypothesis is that two blocks ( $8 \times 8$  pixels) located in the same position of adjacent HS bands are highly similar. There is no universal metric of GOB that is applicable to all HS image in different wavelengths. Moreover, in the experimental results (Zhao et al. 2016; Zhu et al. 2015) indicate that the algorithms get better compression performance at a low bit rate.

Rizzo et al. (2005) and Shen et al. (2017) use a two-stage predictor: one is an inter-band linear predictor, and the other based on least square predictor. The two stage predictor can remove redundancy form two directions, however, the computation cost is also relatively higher in this technique.

DPCM also is an important prediction approach. DPCM include a linear predictor and median filter. The predictor calculates the residual between the value of the current pixel and the predicted pixel. The residual normally has a smaller variance. It results in fewer bits for coding the image. An improved DPCM (Mielikainen and Huang 2012), named as C-DPCM, uses separated spectral clusters. The mean-square error inside each cluster is used to calculate the coefficients. All the pixels used to make the prediction have the same spatial location as the current pixel, and then the difference between the actual targeted value and the predicted value is encoded. This algorithm provides easy mathematical derivation and computational advantage. However, in many practical applications, this still shows an undesirable predictive result.

Paul et al. (2016) proposed a Gaussian mixture-based modelling to predict the succeeding band from current bands. The predicted band is then used as the additional reference band along with the previous band to apply on high efficiency video coding standard (HEVC). Using the number of Gaussian distributions and the initial parameters setting is vital for the result accuracy.

It is worth mentioning that the Consultative Committee for Space Data Systems (CCSDS) have published new compression standard for HS data (Multispectral and Standard 2012). The core predictor in CCSDS called “Fast Lossless (FL)” developed by the NASA Jet Propulsion Laboratory.

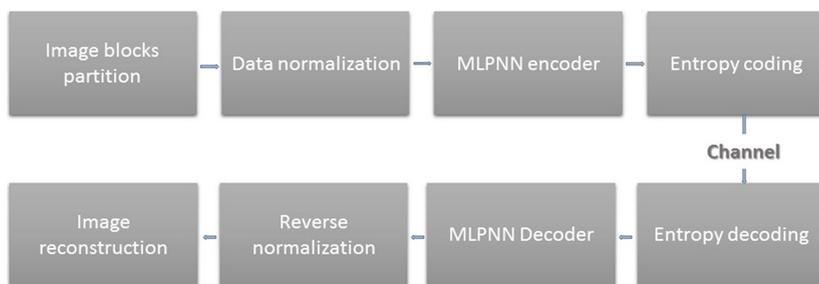
The artificial neural network based coding technique for image compression research is very active and fast development. Especially MLPNN with back propagation training algorithm is one of the most popular NN algorithms and have been used largely in image processing (AL-Allaf 2011; Faris et al. 2016).

In this paper, instead of using MLPNN directly applied to image compression coding, we try innovative MLPNN modelling by its excellent nonlinear approximation capability, finding the hidden relationships between the current band and the succeeding band to achieve high coding efficiency.

Encoding residue can compensate for the error between the predicted band and the previous band. The residue could be near to 0 if the predictor is able to reach an ideal result. On the contrary, bias could be huge. To avoid excessive bias, we use an APM. The procedure of applying quantization step in the encoder and decoder side are both same. In this way, we can save the amount of data that needs to be transmitted.

## 2 Methodology

The multilayer network can be used to approximate almost arbitrary curves if we have enough neurons in the hidden layers (Hagan et al. 1996). The neural network stores the specific information in the weights and biases of the network, which is equivalent to representing the original sample with a smaller data. This is actually a compression process.



**Fig. 2.** A simplified prediction process by the proposed algorithm.

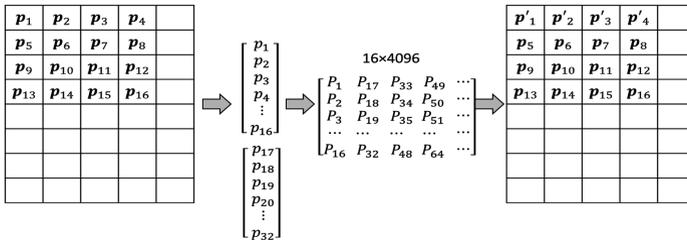
Figure 2 illustrates the prediction process of the proposed schema. The first band is encoded as lossless binary codes then transmitted to the decoder. The MLPNN will be trained on the encoder side and then send weights and biases to the decoder side. For each upcoming band, we need to train once. We have three data space, input data, target data and actual output data. A set of input-output data is generated for each pixel. The prediction process is an iterative process. Considering the first band as input data and the second band as the target output. The predicted band is actual output data of

MLPNN. Then the actual output band will be used as the input band for the next prediction process. That is to say, for the upcoming third band, the decoded 2<sup>nd</sup> band (after compensating the residual with the predicted 2<sup>nd</sup> band with help of decoded bias and weight) is used as the input. The actual third band is the training target. In the decoding site, we reconstruct the predicted band from the bias, weight and the previous band. Then the iterative process will continue until the last band has been trained and reconstructed at the decoder side. Firstly, we'll discuss training data space construction in the next section.

### 2.1 Image Data Pre-process

The data correlation among the neighbouring pixels and the current band  $l$  and the next  $l + 1$  band are the basis for the proposed predicted coding technique. We will use MLPNN learning algorithm based feed forward neural network to predict  $l$  band then encoding the residual. The goal is to find the minimized value of mean square error (MSE) of surrounding bands.

Each band has been resizing to a  $256 \times 256$  matrix and rescales to a range between  $[0, 1]$  that is MPLNN input and output expected values. Then the image was partitioned into  $4 \times 4$  blocks and each block is converted to a  $4^2 \times 1$  vector, e.g. input neuron numbers are 16. Each band converts to a  $16 \times 4096$  matrix as seen Fig. 3.



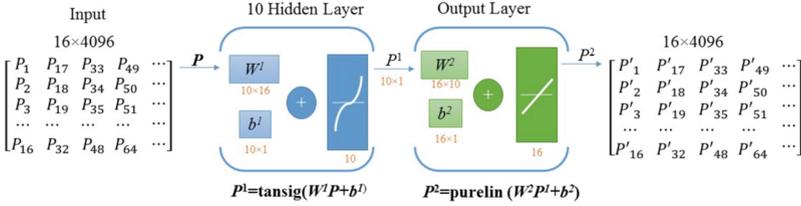
**Fig. 3.** Construction of input space. Input HS images are divided into  $4 \times 4$  blocks of 256 pixels, and the each block is converted to a  $16 \times 1$  vector. The input data is  $l$  band and the output will be predicted  $l + 1$  band

### 2.2 Multilayer Neural Network Architecture

The MLPNN is a 16-10-16 architecture layer. This architecture network can be used as a function approximator. The objective is to find a function that map from the current band (input band) to approximate the reflectance value of the following band (target band). The input data will be a  $16 \times 4096$  matrix of  $l$  band and the output will be predicted  $l + 1$  band. We employed a transfer function tansig as hidden layer and the output layer is the linear function purelin. Using Purelin as output layer data there is no need for it to be normalised, however normalised data can speed up the convergence rate. In the multilayer networks the output of one layer becomes the input to the following layer. The advantage of this structure of network is it can be used as a nonlinear approximator and constrain the outputs of the network between 0 and 1. The

network structure is illustrated at Fig. 4 below. The equations for the hidden layers  $f^1(n)$  and output layer  $f^2(n)$  is given at Eq. (1). Where  $x$  is the net input to a neuron.

$$f^1(x) = \frac{2}{1 + e^{-2x}} - 1 \text{ and } f^2(x) = x \quad (1)$$



**Fig. 4.** MLPNN network structure is a 16-10-16 architecture layer. The tangent sigmoid transfer function is for the Hidden layers and linear function purelin is for the Output layers.

### 2.3 The Error Calculations and Weight Adjustments

The training process of MLPNN is basically tuning the values of the weights and biases of the network to approximate the expected value. The training process first is to propagate the input forward through the network and then propagate the sensitives backward through the network. First we need to choose some initial values for the weights and biases in the range from  $-1$  and  $1$  before training the network. The initial values of weight bias were chosen randomly. Using above Eq. (1) to calculate input and output value of each layer respectively. The mean squared error  $E$  between the target  $P$  and the predictive value  $P'$  is defined as:

$$E = 1/M \sum_{i=1}^M (P_i - P'_i)^2 \quad (2)$$

Where  $M$  is the number of  $P$ . Then the weight vector  $W$  can be updated by

$$\Delta W = W_n - [J^T J + \mu I]^{-1} J^T E \quad (3)$$

Where  $I$  is the identity matrix,  $\mu$  giving as learning rate,  $J$  is Jacobian Matrix can be written as

$$J(x) = \begin{bmatrix} \frac{\partial e_1(x)}{\partial x_1} & \frac{\partial e_1(x)}{\partial x_2} & \dots & \frac{\partial e_1(x)}{\partial x_n} \\ \frac{\partial e_2(x)}{\partial x_1} & \frac{\partial e_2(x)}{\partial x_2} & \dots & \frac{\partial e_2(x)}{\partial x_n} \\ \dots & \dots & \ddots & \vdots \\ \frac{\partial e_N(x)}{\partial x_1} & \frac{\partial e_N(x)}{\partial x_2} & \dots & \frac{\partial e_N(x)}{\partial x_n} \end{bmatrix} \quad (4)$$

Jacobian Matrix is easier to calculate, which don't need to calculate second-order partial derivatives Therefore we choose Levenberg-Marquardt (LM) algorithm as a training function. The LM generally is the fastest training function. It embraces the Gauss-Newton algorithm (GNA) and the method of gradient descent together.

## 2.4 Bands Reconstruction

We present a new idea, only need to encode the weights  $\{W^1\}$   $\{W^2\}$  and biases  $\{b^1\}$   $\{b^2\}$  then send to MLPNN decoder side. MLPNN decoder is the same structure network as encoder side. We only use the previous band  $P_{l-1}$  as an input band, substituted parameters  $\{W^1\}$   $\{b^1\}$  and  $\{W^2\}$   $\{b^2\}$  as shown in Fig. 4. the target band  $P_l$  can be calculated. For each  $16 \times 4096$  band matrix, we got  $n \times 16$   $\{W^1\}$  and  $n \times 16$   $\{W^2\}$  matrix need to be encoded,  $n$  represents the number of hidden layers, e.g. each band with 65536 pixels, MLPNN model using 10 hidden layers will only need to encoded  $10 \times 16$   $\{W^1\}$ ,  $16 \times 10$   $\{W^2\}$  matrix, and  $10 \times 1$   $\{b^1\}$ ,  $16 \times 1$   $\{b^2\}$  vector. We process matrices by mapping minimum and maximum values to  $[0, 255]$  data space then follow a lossless encoding algorithm. The compression ratio is therefore significantly improved. Let's see an example: a size of  $256 \times 256$  image need to encode 65,536 pixels if only encode the weights  $\{W\}$  and biases  $\{b\}$ , we need to encode two  $16 \times 10$  matrixes for weights,  $10 \times 1$  for the first bias and  $16 \times 1$  for the second bias. Their values are from 0 to 255 after mapping. We also need to encode the maximum and minimum values of each matrix for mapping data. The experimental results reveal that the bit requirements for encoding the maximum, minimum, biases and weights are only 1% compared to that of encoding residuals.

## 2.5 Prediction Mechanisms

The predicted band result derived from the target values is further compensated in this step. The ultimate purpose of error-correction learning is bind the relative error between predicted band and the target band. The predicted band is approximated to the targeted band but is not identical, in some cases bias needs to be corrected.

Because the decoder has no original pixel information, refer to Conoscenti et al. (2016), using quantization step sizes in a predictive lossy compression that could reach a near-lossless compression ratio. We also encode quantization step. The process is an adaptive prediction mechanism.

$O_l$  is the compensation value for  $l$ th band,  $O_l = [P_l(1 + \lambda)] - P_l^R$  if the predicted pixel value is smaller than the targeted band pixel, in the case of the predicted pixel is bigger, set  $O_l = [P_l(1 - \lambda)] - P_l^R$ . In the decoder,  $O_l$  will be transmitted in the form of a short integer. This data will be read by the decoder. It is used to recalculate the image data after compensation  $P_l^{R'} = P_l^R + O_l q_l$  is the quantization step size applied to that residual.  $P_l^R$  is the reconstructed pixel.  $\lambda$  represents maximum relative reconstruction error accepted.

### 3 Experiment Result

NASA's AVIRIS HS image dataset is used for our experiments. All of the experimental images have been resized to  $256 \times 256$  for verification purpose. The numbers of neurones of input and output layers are 16. The maximum time of training is 1000. The training goal is set to 0.0001.

The validation of MLPNN model is obtained at the end of each training procedure. MSE gives the difference between observation and simulation.

The performance of training and testing plots the progress in Fig. 5 which indicated that the validation and test curves are very similar. The iteration at 219 Epochs performance reached the minimum MSE.

Figure 6 creates four regression plots for the training set (Train), the validation set (Validation), the test set (Test) and the entire dataset (All). It illustrates the obtained output, target and R-values. The correlation coefficient (R-value) between the outputs and the targets values is well fitted. R-values calculated for evaluating the trained MLPNN model. The reflectance values of pixels are plotted against the targets (circles). The dashed line in each plot represents the perfect result – outputs = targets. The solid line indicates the best linear fit. The R-value related to the training set is very close to 1, which indicates a very good fit. The R-value related to the test set is 0.99636, the training and validation results also show R values that are greater than 0.99.

Figure 7 is a demonstration of the compress results using band 101 ( $l = 100$ ) of the HS image cuprite 1. Predicted  $l + 1$  band is after prediction of the proposed method and the residual between the  $l$  and  $l + 1$  neighbouring bands. The residual images calculate the prediction error between target and predicted  $l + 1$  band, shown in its histogram. The figure are illustrated that the most of residual values are 0 which indicates the MLPNN approximator achieves good results, therefore, result in requiring less number of bits to encode.

Rate/distortion analysis is commonly used for evaluating different encoders which distortion is usually measured in terms of PSNR values. The performance of Fig. 8 shows rate distortion (RD) performance of four HS images Cuprite 1, Cuprite 2,

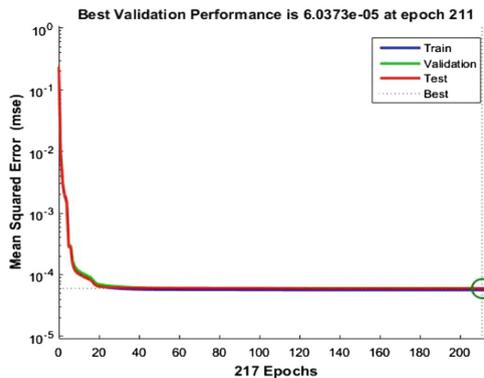


Fig. 5. The validation performance of the MLPNN model

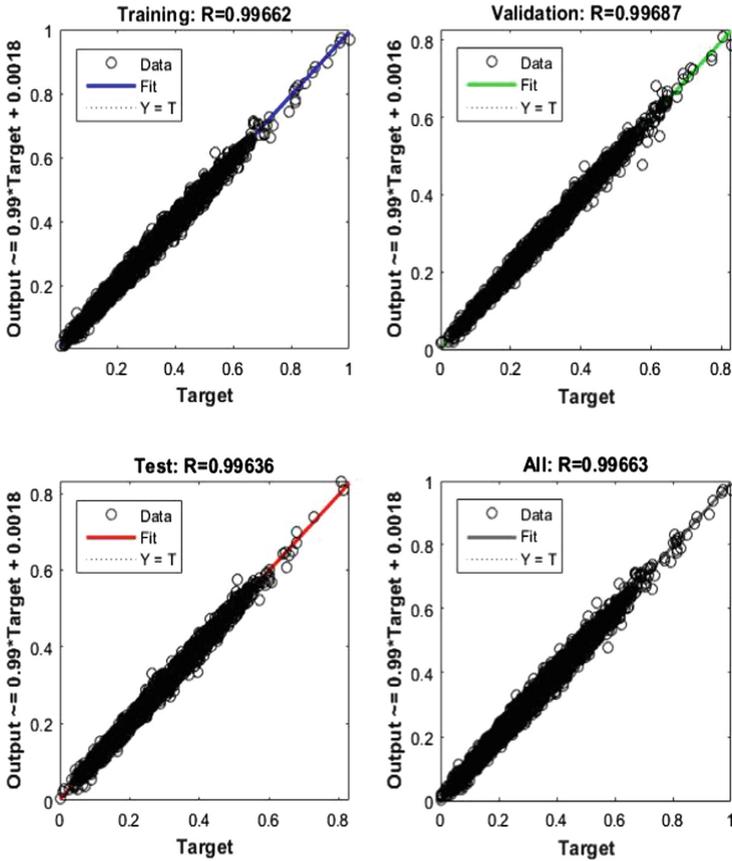
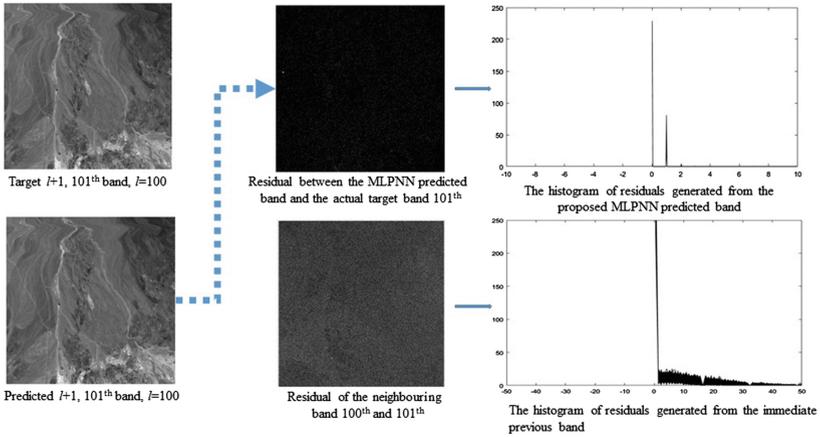
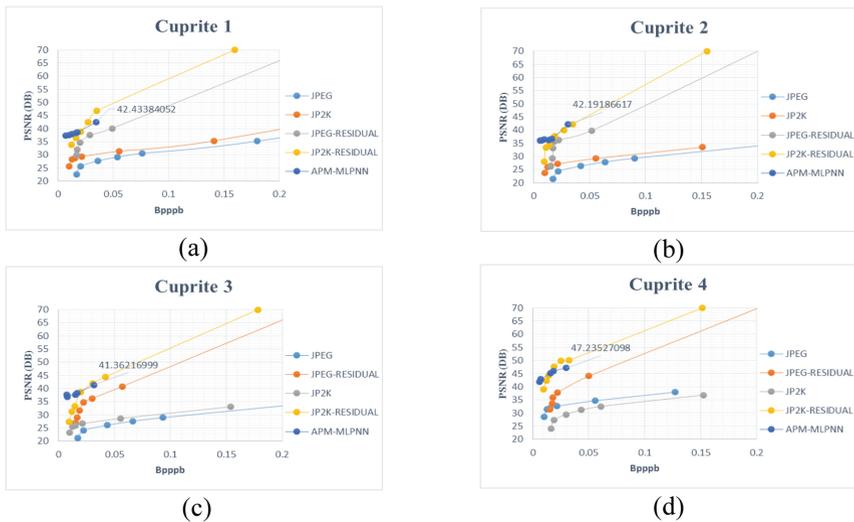


Fig. 6. Performance of the MLPNN model

Cuprite 3 and Cuprite 4 using the proposed MLPNN model, JP2 K, JPEG2 K-residual, JPEG, JPEG-residual encoder techniques. JP2 K is treating each band individually, encoding each band separately by using JPEG 2000 (JP2) image compression standard and coding system. JPEG is the similar method to deal with each band but using the JPEG standard. JP2K-residual and JPEG-residual are to calculate residual of current band and the previous band, then only encode residual instead of encoding each band by JP2 K or JPEG encoder. MLPNN model shows obvious advantages at low bit rate ranges as it only needs to encode weights and biases. APM is to ensure no bias bigger than the maximum accepted reconstruction error. Figure 7 is a demonstration of the compress results using band 101 ( $l = 100$ ) of the HS image cuprite 1. Predicted  $l + 1$  band is after prediction of the proposed method and the residual between the  $l$  and  $l + 1$  neighbouring bands. The residual images calculate the prediction error between target and predicted  $l + 1$  band, shown in its histogram. The figure are illustrated that the most of residual values are 0 which indicates the MLPNN approximator achieves good results, therefore, result in requiring less number of bits to encode.



**Fig. 7.** A demonstration of HS image cuprite 1 compressed result by MLPNN model



**Fig. 8.** Rate distortion performance of four Cuprite HS images using the proposed MLPNN model, JP2K (JPEG 2000), JP2K-residual (JPEG 2000-residual), JPEG, JPEG-residual encoder techniques.

Table 1 illustrates outstanding RD performance at Low bit rates 0.016 bppbb. Both JPEG2 K-residual and JPEG-residual are encoded based on the residuals between the target band  $P_{l+1}$  and the previous band  $P_l$ . Overall, APM-MLPNN model improved the compression rate with very low bit rate costs.

**Table 1.** Rate distortion performance of four Cuprite HS images encoded at low bit rates.

Low bit rates 0.016 bpppb (30 PSNR or more)			
Images	JPEG-residual	JP2K-residual	APM-MLPNN
Cuprite 1	30.03	36.58	38.35
Cuprite 2	33.22	34.44	36.74
Cuprite 3	31.64	33.40	38.25
Cuprite 4	36.02	44.40	45.18

## 4 Conclusion

A hybrid APM with MLPNN model was implemented to predict the succeeding bands. We use the current band as a training data set, and the next band is the training target set in the encoder side. Each band is converted to a  $16 \times 4096$  matrix. Our experimental result achieves better than JPEG and JPEG2 K according to RD performance metrics, PSNR at low bit rates where Bpppb range is from 0 to 0.05. The advance is significant.

The major advantage of this approach is that it only needs to use the binary entropy encoder to encode weights and bias; later in the decoder side the target band can be reconstructed by using transferred weights, bias, residual and the previous band information. Therefore, only a few bits of data need to be transferred. The obtained results were found to be quite satisfactory. The next challenge would be an improvement to reduce the computational time and implementation of deep learning neural networks to further improve the accuracy of the predictive model.

## References

- Aiazzi, B., Baronti, S., Alparone, L.: Lossless compression of hyperspectral images using multiband lookup tables. *IEEE Signal Process. Lett.* **16**, 481–484 (2009)
- AL-Allaf, O.N.A.: Fast backpropagation neural network algorithm for reducing convergence time of BPNN image compression. In: 2011 International Conference on Information Technology and Multimedia (ICIM), pp. 1–6. IEEE (2011)
- Amrani, N., Serra-Sagrìstà, J., Laparra, V., Marcellin, M.W., Malo, J.: Regression wavelet analysis for lossless coding of remote-sensing data. *IEEE Trans. Geosci. Remote Sens.* **54**, 5616–5627 (2016)
- Cheng, G., Han, J.: A survey on object detection in optical remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **117**, 11–28 (2016)
- Cheng, J.-H., Nicolai, B., Sun, D.-W.: Hyperspectral imaging with multivariate analysis for technological parameters prediction and classification of muscle foods: a review. *Meat Sci.* **123**, 182–191 (2017)
- Conoscenti, M., Coppola, R., Magli, E.: Constant SNR, rate control, and entropy coding for predictive lossy hyperspectral image compression. *IEEE Trans. Geosci. Remote Sens.* **54**, 7431–7441 (2016)

- Faris, H., Aljarah, I., Mirjalili, S.: Training feedforward neural networks using multi-verse optimizer for binary classification problems. *Appl. Intell.* **45**, 322–332 (2016)
- Hagan, M.T., Demuth, H.B., Beale, M.H.: *Neural Network Design*, vol. 3632. PWS Publishing Co., Boston (1996)
- Howard, P.G., Vitter, J.S.: New methods for lossless image compression using arithmetic coding. *Inf. Process. Manag.* **28**, 765–779 (1992)
- King, G.R.G., Seldev, C.C., Singh, N.A.: A novel compression technique for compound images using parallel Lempel-Ziv-Welch algorithm. *Appl. Mech. Mater.* **626**, 44 (2014)
- Makki, I., Younes, R., Francis, C., Bianchi, T., Zucchetti, M.: A survey of landmine detection using hyperspectral imaging. *ISPRS J. Photogramm. Remote Sens.* **124**, 40–53 (2017)
- Mielikainen, J., Huang, B.: Lossless compression of hyperspectral images using clustered linear prediction with adaptive prediction length. *IEEE Geosci. Remote Sens. Lett.* **9**, 1118–1121 (2012)
- Mielikainen, J., Toivanen, P.: Lossless compression of hyperspectral images using a quantized index to lookup tables. *IEEE Geosci. Remote Sens. Lett.* **5**, 474–478 (2008)
- Motta, G., Rizzo, F., Storer, J.A. (eds.): *Hyperspectral Data Compression*. Springer Science & Business Media, New York (2006). <https://doi.org/10.1007/0-387-28600-4>
- Multispectral L, Standard HIC: CCSDS 123.0-B-1 Blue Book (2012)
- NASA's Airborne Visible/Infrared Imaging Spectrometer (2017). [https://aviris.jpl.nasa.gov/data/free\\_data.html](https://aviris.jpl.nasa.gov/data/free_data.html)
- Park, B., Lu, R. (eds.): *Hyperspectral Imaging Technology in Food and Agriculture*. FES. Springer, New York (2015). <https://doi.org/10.1007/978-1-4939-2836-1>
- Paul, M., Xiao, R., Gao, J., Bossomaier, T.: Reflectance prediction modelling for residual-based hyperspectral image coding. *PLoS one* **11**, e0161212 (2016)
- Rizzo, F., Carpentieri, B., Motta, G., Storer, J.A.: Low-complexity lossless compression of hyperspectral imagery via linear prediction. *IEEE Signal Process. Lett.* **12**, 138–141 (2005)
- Sasilal, L., Govindan, V.K.: Arithmetic coding-A reliable implementation. *Int. J. Comput. Appl.* **73** (2013)
- Shahriyar, S., Paul, M., Murshed, M., Ali, M.: Lossless hyperspectral image compression using binary tree based decomposition. In: *2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, pp. 1–8. IEEE (2016)
- Shen, H., Pan, W.D., Wu, D.: Predictive lossless compression of regions of interest in hyperspectral images with no-data regions. *IEEE Trans. Geosci. Remote Sens.* **55**, 173–182 (2017)
- Zhao, D., Zhu, S., Wang, F.: Lossy hyperspectral image compression based on intra-band prediction and inter-band fractal encoding. *Comput. Electr. Eng.* **54**, 494–505 (2016)
- Zhu, S., Zhao, D., Wang, F.: Hybrid prediction and fractal hyperspectral image compression. *Math. Probl. Eng.* **2015**, 10 (2015)