Data Aided Sensing with Variational Autoencoder

Muhammad Awais, *Gradute Student Member, IEEE*, Jinho Choi, *Fellow, IEEE*, Jihong Park, *Senior Member, IEEE* and Yun Hee Kim, *Senior Member, IEEE*

***Abstract*—In this paper, we propose an intelligent data-aided sensing method exploiting variational auto-encoder (VAE) for wireless sensor networks. A set of measurements of all sensors is analyzed to identify the underlying cause. For this purpose, we employ β-VAE to find the cause with a partial observation of the measurement vector and utilize the inferred cause for communication efficient data collection in uploading the data. The resulting approach can be used when collecting all measurements would be expensive, or excessively time consuming.**

***Index Terms*—Data-aided sensing, Deep learning, Clustering, Semantic sensing, Internet of things, Variational autoencoder**

I. Introduction

The sixth generation (6G) networks are expected to support various Internet-of-Things (IoT) applications for smart cities, autonomous industries, sustainable environments, and so forth [1]. In those applications, there are numerous connected sensors and IoT devices that can provide various operational states and environmental information. While measurements can be collected from sensors and devices through IoT platforms, their analysis and interpretation are other tasks that require additional effort along with domain knowledge. Most studies have considered sensing and analysis as separate tasks [2] so that the measurements from sensors are collected without understanding their meaning in general.

Semantic communication has recently emerged as a key technology in implementing the 6G networks [3]. The goal of semantic communication is to convey the meaning of a message rather than simply transmit a bit sequence over a noisy channel reliably. The main concept of semantic communication can be applied to sensing for efficient data collection and analysis in a given application. To understand the meaning of sensed data, analysis is required using various techniques. Through analysis, it is also possible to improve the data collection or reduce the communication cost. In particular, in [4]–[6], the notion of data-aided sensing (DAS) is proposed to reduce the communication cost by utilizing the dataset collected earlier to find the sensors with more informative measurements than others, based on statistical knowledge of sensors' measurements.

In this paper, we propose a type of semantic sensing, a process of extracting meaningful information from sensor data and using it for efficient data collection in the later stage. We regard each sensor's measurement as a pixel of a given image while considering the measurements of all sensors, which is referred to as the complete measurement vector, as an image, which is a result of a cause. We address β-variational auto-encoder (VAE) [7], [8] followed by a clustering method to find a semantic/cause with a partial observation on the measurement vector and to facilitate DAS based on the found semantic/cause as well as to estimate the unseen part of the measurement vector.

The semantic/cause obtained is also used for DAS to select the next sensor with a more informative measurement to reach to a reliable detection level faster with a smaller size of partial observation. Thus, the resulting method can reduce the communication cost significantly.

II. System Model with Data-Aided Sensing

Suppose a IoT network, where a BS communicates with M sensor nodes. The measurement of M sensor nodes is denoted by random vector $\mathbf{X} = [X_1 \ X_2 \ \cdots \ X_M]^T \in \mathbb{R}^M$ of which the instance is expressed as $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_M]^T \in \mathbb{R}^M$. We assume that a measurement is the result of a certain cause or factor listed in the set $\mathcal{N}$, where the number of causes is much smaller than the size of the measurement vectors as $N(= |\mathcal{N}|) \ll M$. For instance, a set of MNIST handwritten digit images, where the values of pixels of each image can be seen as the measurements of M sensors. In this case, the cause $\mathcal{C}_n = n-1$ results in handwritten digit images of number $n-1$ for $n \in \mathcal{N} \triangleq \{1, 2, \cdots, N\}$.

To find the cause of a complete measurement, the BS can send a query to a specific node m) and receive its response x_m at a time. Thus, in order to have a complete measurement, there should be M rounds. Let $m(l)$ denote the index of the node that uploads its local measurement at round l for a given cause $\mathcal{C}_n$. The set of local measurements that are available at the BS after l rounds is given by

$$\mathcal{X}(l) = \{x_m, \ m \in \mathcal{M}(l)\}, \quad (1)$$

where $\mathcal{M}(l) = \{m(1), \cdots, m(l)\}$.

In this paper, we wish to understand semantic/cause $\mathcal{C}_n$ of the measurement with partial observation $\mathcal{X}(l)$ in a limited amount of time or to minimize the communication cost, and possibly generate an estimate of the complete measurement. In

M. Awais and Y.H. Kim are with the Department of Electronics and Information Convergence Engineering, Kyung Hee University, Yongin 17104, Korea (e-mail: {mawais, yheekim}@khu.ac.kr).

J. Choi and J. Park are with School of Information Technology, Deakin University, Melbourne, Australia (e-mail: {jinho.choi, jihong.park}@deakin.edu.au).

This research was supported by the National Research Foundation of Korea (NRF) under Grant NRF-2021R1A2C1005869 and by the Institute of Information & Communications Technology Planning & Evaluation (IITP) under the Information Technology Research Center (ITRC) support program IITP-2023-2021-0-02046, with funding from the Ministry of Science and ICT (MSIT), Korea (*Corresponding author: Yun Hee Kim*).

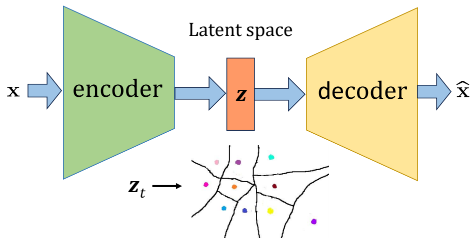


Fig. 1: VAE model and an illustration of the latent vectors as causes resulting in $\mathbf{x}$.

addition, before the cause of $\mathcal{C}_n$ of measurements is detected at a certain level of accuracy, the BS may perform communication cost-efficient data collection [5] based on semantic sensing. For instance, we apply DAS that decides the next node to upload its local information for a certain objective function, $\Theta(\cdot)$, as

$$m(l+1) = \arg \min_{m \in \mathcal{M}^c(l)} \Theta(X_m | \mathcal{X}(l)), \tag{2}$$

where the selected node $m(l+1)$ in each round depends on the accumulated measurements from the previous rounds, $\mathcal{X}(l)$.

III. Application of VAE for DAS

In this section, we propose to employ β-VAE to understand semantic/cause $\mathcal{C}_n$ of sensors' measurements and predict the complete measurement with partial observation $\mathcal{X}(l)$.

A. β-VAE for Intelligent Sensing

The structure of the VAE [7] is depicted in Fig. 1, where the dimension of the latent space is K ($\ll M$). The VAE encoder $e_\phi(\mathbf{x})$ parameterized by ϕ maps the input vector $\mathbf{x}$ to $\mathbf{z}$ to have conditional distribution $q_\phi(\mathbf{z}|\mathbf{x})$. The VAE decoder $d_\theta(\mathbf{z})$ parameterized by θ maps from the latent space $\mathbf{z}$ to the input space $\hat{\mathbf{x}}$ to have the conditional distribution $p_\theta(\mathbf{x}|\mathbf{z})$.

For efficient clustering in the latent domain and reproducibility of a complete measurement, we employ the β-VAE loss function as [9]

$$\mathcal{L}(\theta, \phi; \mathbf{x}, \mathbf{z}, \beta) = \mathcal{L}_{Rec} + \beta \mathcal{L}_{KL}, \tag{3}$$

where $\mathcal{L}_{Rec} = -E_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}|\mathbf{z})]$ is the reconstruction term controlling the quality of the generated data $\hat{\mathbf{x}}$ from the latent representation, $\mathcal{L}_{KL} = D_{KL}(q_\phi(\mathbf{z}|\mathbf{x}) || p(\mathbf{z}))$ is the regularization term minimizing the Kullback–Leibler (KL) divergence between the approximate posterior $q_\phi(\mathbf{z}|\mathbf{x})$ and prior $p(\mathbf{z})$, and $\beta > 0$ is the hyperparameter balancing the above two terms.

B. Training and Clustering

Suppose that a training dataset $\mathcal{T} = \{\mathbf{x}_t\}_{t=1}^{|\mathcal{T}|}$ of complete measurement vectors is given. For given $\mathcal{T}$ and β, the VAE is trained to minimize the loss (3). Once the VAE is trained, we obtain the latent vectors $\mathcal{Z} = \{\mathbf{z}_t, t = 1, 2, \cdots, |\mathcal{T}|\}$ of the training dataset $\mathcal{T}$ with the trained VAE. The latent vectors in $\mathcal{Z}$ are then clustered with the k-means clustering algorithm [10] as $\mathbf{z}_t \in \mathcal{Z}_{(n)}$ for $n \in \mathcal{N}$, which also provides the centroid $\bar{\mathbf{z}}_{(n)}$ of each cluster $\mathcal{Z}_{(n)}$. We also compute the covariance matrix $\mathbf{\Sigma}_{(n)}$ of the latent vectors in cluster $\mathcal{Z}_n$ as

$$\mathbf{\Sigma}_{(n)} = \tfrac{1}{|\mathcal{Z}_{(n)}|} \sum_{\mathbf{z}_t \in \mathcal{Z}_{(n)}} (\mathbf{z}_t - \bar{\mathbf{z}}_{(n)})(\mathbf{z}_t - \bar{\mathbf{z}}_{(n)})^T. \tag{4}$$

The centroid $\bar{\mathbf{z}}_{(n)}$ and covariance matrix $\mathcal{Z}_{(n)}$ of each cluster $\mathcal{Z}_{(n)}$ are stored at the BS for semantic sensing.

C. Semantic Sensing with trained VAE and DAS

The BS operates with semantic sensing once the VAE is trained and the clustering information is obtained. Assume that a sample instance $\mathbf{x}$ of $\mathbf{X}$ corresponding to a particular cause $\mathcal{C}_n$ is measured at the sensors. After the lth round of data uploading, the BS aims to identify the cause $\mathcal{C}_n$ with partial observation $\mathcal{X}(l)$ on $\mathbf{x}$ available at the BS.

We propose to infer the cause in the latent space by constructing the input for the VAE encoder as

$$\tilde{\mathbf{x}}(l) = \mathcal{X}(l) \cup \{[\bar{\mathbf{x}}_A]_{i \in \mathcal{M}^c(l)}\}, \tag{5}$$

where $\bar{\mathbf{x}}_A = \sum_{n=1}^{N} \bar{\mathbf{x}}_{(n)} \hat{P}(\mathcal{C}_n)$ with $\bar{\mathbf{x}}_{(n)} = \mathrm{dec}_\theta(\bar{\mathbf{z}}_{\mathrm{(n)}})$ is the average of the representative measurement vectors of all causes and $\hat{P}(\mathcal{C}_n) = |\mathcal{Z}_{(n)}|/|\mathcal{T}|$. The augmented input is then mapped to a latent vector by the VAE encoder as $\boldsymbol{\mu}_A(l) = \mathrm{enc}_\phi(\tilde{\mathbf{x}}_A(l))$. We then apply the MAP criterion for cause selection by assuming the Gaussian distribution of the latent vectors in each cluster $\mathcal{Z}_{(n)}$ as

$$\hat{n}_{\mathrm{map}}(l) = \arg \max_{n \in \mathcal{N}} e^{-\Phi_n(\boldsymbol{\mu}_A(l))} = \arg \min_{n \in \mathcal{N}} \Phi_n(\boldsymbol{\mu}_A(l)), \tag{6}$$

where

$$\Phi_n(\mathbf{z}) = \frac{1}{2}(\mathbf{z} - \bar{\mathbf{z}}_{(n)})^T \mathbf{\Sigma}_{(n)}^{-1} (\mathbf{z} - \bar{\mathbf{z}}_{(n)}) + \ln\left(\frac{|\mathbf{\Sigma}_{(n)}|}{\hat{P}(\mathcal{C}_n)}\right). \tag{7}$$

The MAP detection can be further simplified into the Euclidean distance detection.

With the cause detected, we further apply the DAS for communication cost-effective data selection. By assuming the Gaussian distribution on the measurement vector, we may apply the mean square error (MSE) since we can reduce the MSE by selecting the sensor with the largest MSE. In semantic sensing, the objective function for DAS may depend on the cause of a measurement. By using the conditional MSE as

$$\mathrm{MSE}(X_m | \mathcal{C}_n) = \mathbb{E}\left[|X_m - \mathbb{E}[X_m \,|\, \mathcal{C}_n]|^2 \,\middle|\, \mathcal{C}_n \right], \tag{8}$$

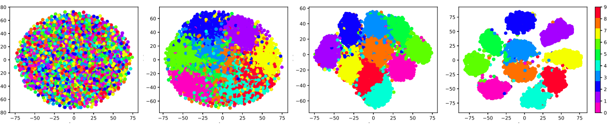
we adopt the causewise DAS (C-DAS) as

$$\Theta(X_m | \mathcal{X}(l)) = -\mathrm{MSE}(X_m | \mathcal{C}_{\hat{n}}) \tag{9}$$

by using the cause $\mathcal{C}_{\hat{n}}$ detected or the weighted C-DAS (WC-DAS) incorporating the effect of erroneous detection as

$$\Theta(X_m | \mathcal{X}(l)) = -\sum_{n \in \mathcal{N}} \mathrm{MSE}(X_m | \mathcal{C}_n) \tilde{P}_n \tag{10}$$

with $\tilde{P}_n = \frac{\exp(-\Phi_n(\boldsymbol{\mu}_A(l)))}{\sum_{j \in \mathcal{N}} \exp(-\Phi_j(\boldsymbol{\mu}_A(l)))}$.



(a) $\beta = 10$ (b) $\beta = 1$ (c) $\beta = 0.1$ (d) $\beta = 0.01$

Fig. 2: t-SNE representation of the latent space vectors $\mathbf{z}$ for the clean MNIST dataset.

IV. Experimental Setup and Results

We investigate the feasibility of VAE-aided semantic sensing with the MNIST dataset. Each pixel of a 28×28 MNIST image represents a measurement of a sensor for $M = 784$ sensors. The clean and noisy MNIST datasets are employed, where zero-mean Gaussian noises with variance $\sigma_x^2 = 0.3$ are added to the images for the noisy MNIST dataset. We used 60k and 10k data for training and testing, respectively for each dataset. The VAE with convolutional neural networks and $K = 20$ for the dimension of latent space is employed by following the network model [9]. The VAE is trained by employing the Adam optimizer with a learning rate 10^{-3} and a batch size 32. The initial cluster centroids for k-means clustering are chosen randomly from the latent vectors.

We first find an appropriate choice of β in β-VAE for the MINST dataset by examining the clustering behaviours of the latent vectors in Fig. 2 when $\beta = 10$, 1, 0.1, and 0.01, respectively. The t-distributed stochastic neighbour embedding (t-SNE) is applied to visualize the latent space vectors in two dimensions by assigning different colours for $\mathbf{z}$ corresponding to different MINIST digits. We choose $\beta = 0.01$ in training the VAE for appropriate clustering since the original VAE with $\beta = 1$ fails to classify the latent vectors.

The average cause detection probability is compared as the percentage of active sensors uploading the measurement increases for the clean MNIST dataset in Fig. 3 and for the noisy MNIST dataset in Fig. 4. The MAP and ED detection combined with WC-DAS (10), C-DAS (9), and random node selection (RAND) are compared. The results show that WC-DAS and C-DAS with MAP improve the average detection probability significantly when compared with RAND. A slight gain of WC-DAS over C-DAS is observed when the correct detection error probability is low but it is almost negligible. A loss of ED over MAP is also observed and the loss does not diminish even with the complete measurement. Due to the noises, the average cause detection probability of the noisy dataset is limited to around 0.8 in Fig. 4.

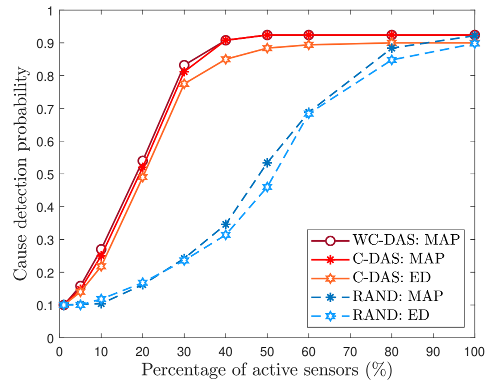


Fig. 3: Average cause detection probability for the clean MNIST dataset when $\beta = 0.01$ for the VAE.

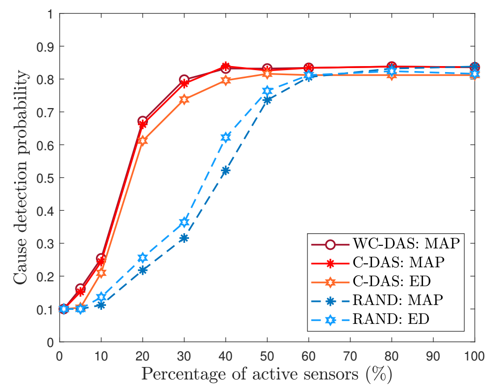


Fig. 4: Average cause detection probability for the noisy MNIST dataset when $\beta = 0.01$ for the VAE.

V. Concluding Remarks

This paper has proposed VAE-aided intelligent sensing which detects a cause of sensors' measurements with their partial observation. By exploiting a β-VAE and k-means clustering in the latent space, the cause is detected under the MAP criterion in the latent space and is used for DAS. The proposed method was verified with the MNIST dataset having 10 causes by regarding each pixel of an image as a measurement of a sensor. The cause detection probability with a partial observation is improved significantly with DAS.

References

[1] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, D. Niyato, O. Dobre, and H. V. Poor, “6G internet of things: A comprehensive survey,” *IEEE Internet Things J.*, vol. 9, pp. 359–383, Jan. 2022.

[2] Y. Kuo, C. Li, J. Jhang, and S. Lin, “Design of a wireless sensor network-based IoT platform for wide area and heterogeneous applications,” *IEEE Sensors J.*, vol. 18, pp. 5187–5197, June 2018.

[3] E. C. Strinati and S. Barbarossa, “6G networks: Beyond shannon towards semantic and goal-oriented communications,” *Comput. Netw.*, vol. 190, p. 107930, May 2021.

[4] J. Choi, “A cross-layer approach to data-aided sensing using compressive random access,” *IEEE Internet Things J.*, vol. 6, pp. 7093–7102, Apr. 2019.

[5] J. Choi, “Gaussian data-aided sensing with multichannel random access and model selection,” *IEEE Internet Things J.*, vol. 7, pp. 2412–2420, Mar. 2020.

[6] J. Choi, “Data-aided sensing for Gaussian process regression in IoT systems,” *IEEE Internet Things J.*, vol. 8, pp. 7717–7726, Sep. 2021.

[7] D. P. Kingma and M. Welling, “Auto-encoding variational bayes,” *arXiv preprint arXiv:1312.6114*, 2014.

[8] D. P. Kingma and M. Welling, “An introduction to variational autoencoders,” *Foundations and Trends Mach. Learn.*, vol. 12, no. 4, pp. 307–392, Nov. 2019.

[9] I. Higgins, L. Matthey, A. Pal, C. Burgess, X. Glorot, M. Botvinick, S. Mohamed, and A. Lerchner, “beta-VAE: Learning basic visual concepts with a constrained variational framework,” in *Proc. Int. Conf. Learn. Represent. (ICLR), Toulon, France*, Apr. 2017.

[10] C. M. Bishop and N. M. Nasrabadi, *Pattern Recognition and Machine Learning*. Berlin, Heidelberg: Springer-Verlag, 2006.