**Encoder and Decoder Usage in Real Applications**

Michael Stewart

Colorado State University Global

CSC580-1: Applying Machine Learning and Neural Networks - Capstone

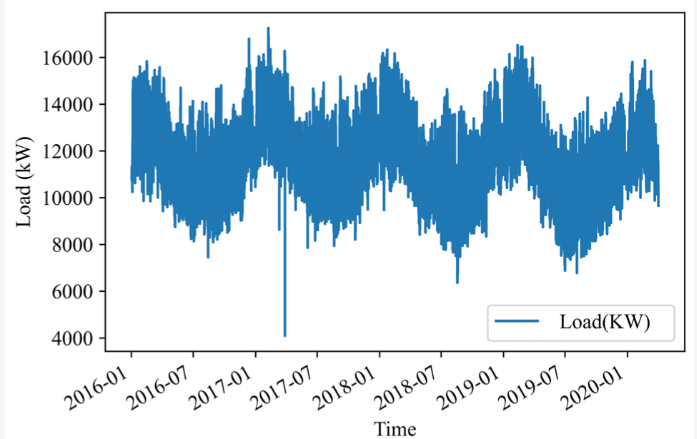
Professor Isaac Gang

November 3, 2024

**Encoder and Decoder Usage in Real Applications**

Encoder and decoder systems in machine learning and artificial intelligence are useful tools for processing sequential data. Most of the sequential data processed by these systems are natural language processing and image processing. This makes encoder/decoder systems useful for a large amount of real-life applications. Since machine learning is a recent development, we have only had a short time to explore how this technology can be applied to our lives. The examples talked about in this paper will include an electric load forecasting system with consumption and meteorological data as inputs, a machine vision system tasked with pavement crack detection, a manufacturing surface inspection system, and a medical image segmentation system.

**Short-Term Campus Load Forecasting**

****

In this system, the goal is to forecast the electrical load of the Memorial University of Newfoundland. To do this, the university team took electric consumption data and meteorological data as an input to an encoder. Then the model will encode this information using an attention method. The developers justify their use of the attention method “The motivation for using the attention mechanism stems from the underperformance of the basic encoder-decoder model for long sequences. This is caused by the use of a fixed-length context vector.” (Ahmed, Jamil, & Khan, 2024). The model then uses the processed and encoded data to provide an electrical load prediction for the next day. The results showed that the model using attention performed very well compared to the decoder model without attention. One of the interesting things about this model is that it will learn from previous data on electrical consumption and give a good guess about power consumption.

**Pavement Crack Detection**

In this example, there is a huge problem with the quality of roads and deterioration over time. Prioritizing which road’s surface should be fixed first is an important part of the infrastructure. The model proposed will detect cracks using an unmanned aerial vehicle and an autoencoder convolutional neural network (Fakhri et al., 2023). This is a great method to use for machine vision and crack detection because encoder-decoder architecture is good at converting images into data the model can understand and perform analysis. The results of the paper show that the encoder-decoder system performed well because of its ability to contextualize different markings in the photos. Based on its sequential nature, the system is able to detect cracks in surfaces with an accuracy of 98%.

**Automatic Surface Detection in Manufacturing Quality Control**

This example is similar to the previous one because it uses an encoding and decoding system to analyze images with imperfections in a surface. However, this system is used to identify specific markings and imperfections with manufacturing product surfaces. To do this, the model needs to keep specific details in its resolution so it can classify them. It uses a multi-dimensional feature extraction-based deep encoder-decoder network to solve this problem (Uzen et al., 2023). One interesting feature of this system is that they used separate datasets in order to evaluate the accuracy with different training. The group was able to achieve accuracies of 80% and 56% with each of the datasets used for training. One of the things I would try next for this team is to train the model on both datasets or a mixture of them.

**Medical Image Segmentation**

Finally, the last example of a real-life application used for encoder-decoder models is image segmentation in the medical field. The encoder-decoder system used in this example is used for a completely different reason than the other ones I have mentioned so far. This example uses an encoder to sort through the photos, lower the images’ resolution, and select potential segmentations in the photo (Yang et al., 2024). The encoder then feeds the decoder the segmented data from an image and the decoder has the job of identifying the segmented photo, expanding the resolution, and converting it back into an image that humans can understand. Because the encoder and decoder are separate structures in this system, the system used for the analysis of the segmented photo can only focus on the segmented image. It will not be thrown off by complicated images that have features interfering with each other. In conclusion, encoders and decoders in this system separate the data whole data and do not use context which is an interesting implementation of this system since it usually uses context in machine vision.

**References**

Ahmed, Z., Jamil, M., & Khan, A. A. (2024). Short-Term Campus Load Forecasting Using CNN-Based Encoder–Decoder Network with Attention. Energies (19961073), 17(17), 4457. <https://doi.org/10.3390/en17174457>

Fakhri, S. A., Satari Abrovi, M., Zakeri, H., Safdarinezhad, A., & Fakhri, S. A. (2023). Pavement crack detection through a deep-learned asymmetric encoder-decoder convolutional neural network. International Journal of Pavement Engineering, 24(1), 1–12. <https://doi.org/10.1080/10298436.2023.2255359>

Ganegedara, T. (2022). *TensorFlow in action*. Simon and Schuster.

Uzen, H., Turkoglu, M., & Hanbay, D. (2023). Multi-dimensional feature extraction-based deep encoder–decoder network for automatic surface defect detection. Neural Computing & Applications, 35(4), 3263–3282. <https://doi.org/10.1007/s00521-022-07885-z>

Yang, Z., Sun, X., Yang, Y., & Wu, X. (2024). MEDU-Net+: a novel improved U-Net based on multi-scale encoder-decoder for medical image segmentation. KSII Transactions on Internet and Information Systems, 18(7), 1706. <https://doi.org/10.3837/tiis.2024.07.001>