World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:17, No:05, 2023

Categorical Metadata Encoding Schemes for Arteriovenous Fistula Blood Flow Sound Classification: Scaling Numerical Representations Leads to Improved Performance

Authors : George Zhou, Yunchan Chen, Candace Chien

Abstract : Kidney replacement therapy is the current standard of care for end-stage renal diseases. In-center or home hemodialysis remains an integral component of the therapeutic regimen. Arteriovenous fistulas (AVF) make up the vascular circuit through which blood is filtered and returned. Naturally, AVF patency determines whether adequate clearance and filtration can be achieved and directly influences clinical outcomes. Our aim was to build a deep learning model for automated AVF stenosis screening based on the sound of blood flow through the AVF. A total of 311 patients with AVF were enrolled in this study. Blood flow sounds were collected using a digital stethoscope. For each patient, blood flow sounds were collected at 6 different locations along the patient's AVF. The 6 locations are artery, anastomosis, distal vein, middle vein, proximal vein, and venous arch. A total of 1866 sounds were collected. The blood flow sounds are labeled as "patent" (normal) or "stenotic" (abnormal). The labels are validated from concurrent ultrasound. Our dataset included 1527 "patent" and 339 "stenotic" sounds. We show that blood flow sounds vary significantly along the AVF. For example, the blood flow sound is loudest at the anastomosis site and softest at the cephalic arch. Contextualizing the sound with location metadata significantly improves classification performance. How to encode and incorporate categorical metadata is an active area of research1. Herein, we study ordinal (i.e., integer) encoding schemes. The numerical representation is concatenated to the flattened feature vector. We train a vision transformer (ViT) on spectrogram image representations of the sound and demonstrate that using scalar multiples of our integer encodings improves classification performance. Models are evaluated using a 10-fold cross-validation procedure. The baseline performance of our ViT without any location metadata achieves an AuROC and AuPRC of 0.68 ± 0.05 and 0.28 ± 0.09, respectively. Using the following encodings of Artery:0; Arch: 1; Proximal: 2; Middle: 3; Distal 4: Anastomosis: 5, the ViT achieves an AuROC and AuPRC of 0.69 ± 0.06 and 0.30 ± 0.10, respectively. Using the following encodings of Artery:0; Arch: 10; Proximal: 20; Middle: 30; Distal 40: Anastomosis: 50, the ViT achieves an AuROC and AuPRC of 0.74 ± 0.06 and 0.38 ± 0.10, respectively. Using the following encodings of Artery:0; Arch: 100; Proximal: 200; Middle: 300; Distal 400: Anastomosis: 500, the ViT achieves an AuROC and AuPRC of 0.78 ± 0.06 and 0.43 ± 0.11. respectively. Interestingly, we see that using increasing scalar multiples of our integer encoding scheme (i.e., encoding "venous arch" as 1,10,100) results in progressively improved performance. In theory, the integer values do not matter since we are optimizing the same loss function; the model can learn to increase or decrease the weights associated with location encodings and converge on the same solution. However, in the setting of limited data and computation resources, increasing the importance at initialization either leads to faster convergence or helps the model escape a local minimum.

Keywords : arteriovenous fistula, blood flow sounds, metadata encoding, deep learning
Conference Title : ICASSP 2023 : International Conference on Acoustics, Speech and Signal Processing
Conference Location : New York, United States
Conference Dates : May 15-16, 2023

Open Science Index, Electrical and Computer Engineering Vol:17, No:05, 2023 publications.waset.org/abstracts/163552.pdf