**Meta-Learning Wavelet Neural Operators for Personalized Generative Therapy**

The quest to objectively quantify real-time emotional responses has perennially challenged researchers—from Spielberger’s subjective State Trait Anxiety Inventory (STAI) to Paul Ekman’s pioneering, yet unfulfilled, attempts to use decision trees for predicting emotions from physiological signals. These foundational efforts, though critical, have revealed a stark absence of objective, universal methodologies capable of quantifying the spectrum of emotional and anxiety states. This deficiency impedes the advancement of emotionally intelligent systems that are tailored to individual mental states, precise mental health diagnostics, and personalized telemental healthcare.

Recent machine learning advancements have carved new pathways for linking physiological signals to emotional states. However, these methodologies typically demand extensive data sets and uniform sensor usage across those datasets—a notable challenge given the diminutive and heterogeneous nature of existing emotional databases like WESAD and EMOGNITION. Our research confronts these hurdles directly by introducing a ground-breaking zero-shot wavelet neural operator for signal encoding. This operator maps physiological signals onto a common manifold, enabling the amalgamation of disparate datasets into a single model and thus significantly broadening the trainable domain.

Central to our approach is the innovative wavelet neural operator technique, designed specifically to tackle the complex challenges of encoding time-series physiological data. By operating within the wavelet domain, this technique ensures high-resolution feature capture across both time and frequency dimensions, substantially enhancing the model's interpretability and resilience to noise. Unlike traditional deep learning approaches, our method offers a more refined analysis of physiological signals, leading to greatly improved accuracy in emotion predictions.

To validate our model, we present our custom-made wearable patch that records emotionally relevant physiological signals. We introduce this dataset as well as present a new integrated wearable device for future emotion research that is compatible with virtual reality goggles. The intrinsic softness, thickness, and gas permeability of skin-interfaced electronics hold promise for long-term and continuous high-fidelity monitoring of physiological states. Recently, there has been many exciting progresses in the fabrication of ultrathin and soft wearables, as evidenced by elastomeric nanofiber mats1 and elastic conductor with microcracked structures and semiconductors. While these devices exhibited seamless contact with skin, the fabrication typically involves electrospinning, iterative spin coating, and thermal evaporation, which are laborious and time-consuming. Moreover, the manufacture of large patches to cover the head or face presents challenges due to limitations in dimensions imposed by instruments such as spin-coaters and oxygen plasma cleaners. While the float assembly method shows promise, achieving in-situ generation of microscale porosity remains fundamentally challenging.

Our approach was validated through generative virtual reality, music, heat, and voice therapy sessions. Each therapy is purely generated from the output of our model describing the emotional profile of the patient. In our early tests, we have shown that we can successfully modulated and reduce negative emotions and anxiety. We validate our model’s prediction against the current gold standard STAI questionnaires. This validation not only supports Ekman’s theoretical framework linking physiological and psychological domains but also sets a new benchmark for personalized therapeutic interventions.

This research introduces a robust framework for the objective and accurate prediction of emotional states from physiological data, marking a significant leap forward in the fields of affective computing and personalized medicine. Here, we present an ultrathin porous electronic skin based on air/water interfacial assembly that achieves substantial improvement in unobtrusiveness, comfort, and intimate contact with skin. This approach enables rapid (in minutes), facile and large-area (>200 cm2) fabrication of ultrathin patches (~1 µm) and phase-separated porous elastic sensors (~6 µm). Our findings promise to revolutionize how we understand and interact with emotional dynamics in both clinical and everyday settings.

More information:

The capability to objectively quantify real-time emotional responses has always been an elusive area of research from the development of the subjective State Trait Anxiety Inventory (STAI) exam by Spielberger to Paul Ekman’s unrealized approach to use decision trees to predict emotions from physiological parameters. In lieu of the difficulties, it remains imperative to find an objective universal approach for quantifying different emotions and anxiety levels for emotionally intelligent systems to adapt to a user’s state, provide a new objective approach for mental health diagnosis, and provide personalized real-time telemental healthcare. Nevertheless, the prevailing opinion amongst medical professionals is that emotional responses are linked (and even require) physiological responses. Modern advances in machine learning offer a new approach to link physiological signals to the emotional state. The main problem with utilizing recent machine learning architectures (such as transformers) is that they require a huge amount of data, a set number of features, and the same sensors across all data points. Unfortunately, current emotion databases such as the WESAD and EMOGNITION datasets are small (< 500 points), use different sensors (ECG, EEG, etc), and a different number of physiological parameters. To tackle this challenge, we have developed a zero-shot approach for signal encoding that introduces the world’s first wavelet neural operator to mathematically map physiological signals onto a common manifold domain, thus allowing us to feed multiple datasets through a single model, drastically increasing the trainable domain. Through this meta-learning technique, we showed high accuracy in predicting a user’s state purely from their physiological parameters, affirming Dr. Ekman’s direction in linking the physiological to the psychological spaces. Furthermore, we validated our model by selectively reducing negative emotions and anxiety levels in patients through generative virtual reality therapy, based on the predicted emotional state from our model. This novel approach not only pioneers personalized therapeutic interventions but also opens new avenues for understanding emotional dynamics in virtual environments.

We introduce a pre-trained pseudo-deterministic universal signal encoder that is capable of out-of-domain encoding, mapping an arbitrary number of n-length time series signals to a meaningful physiological-relevant latent space. By conserving the spatial and temporal properties of the initial sensor data, the encoding space becomes interpretable and hence provides clinically relevant information about how the model utilizes each sensor’s information. This contrasts the recent general signal encoders in the literature that utilize transformer-based architectures, which requires a huge dataset to train as well as fine-tuning for out-of-domain examples.

Our zero-shot domain-agnostic physiological encoder is trained on the world’s first wavelet neural operator: a potentially universal method for learning information from complex time-series physiological data with both good frequency and time resolution. By constraining the analysis to neural operators, we preserve the spatial and temporal information by learning a mapping between the time and wavelet domain. By learning the information in the wavelet domain, we maintain self-attention to all components of the signal, split by high and low frequency decompositions. After learning the information in the wavelet domain, we can easily map the solution back into a physiologically relevant time-series signal that can be further subjective to pattern recognition and complex interpretations. Furthermore, as the Jacobian between the time and wavelet domain is greater than one, small prediction errors in the wavelet domain are reduced once mapped back to the time domain, acting as an inherent denoiser between training instances. This allows he model to visually converge to the same encoded signal in the time domain, despite small noticeable differences in the wavelet domain.