**Scientific Abstract**

**Title: Generalization in Deep Learning: The Interplay of Adaptive Strategies for Normalization, Optimization, and Regularization**

Generalization lies at the heart of deep learning, driving its profound successes and transformative potential. This fundamental concept permeates numerous domains, from computer vision and natural language processing to autonomous driving and healthcare. Generalization is the ability to extrapolate knowledge acquired during training to new, unseen data, without overfitting to the training data’s characteristics. Among other things, deep learning generalization depends on core elements such as data normalization, model optimization, and regularization. Normalization techniques have become essential in mitigating the impact of data variability. By standardizing input data, normalization methods facilitate the extraction of meaningful features, thereby fostering generalization. Optimization algorithms, which are divided into adaptive and non-adaptive methods, significantly influence the learning trajectory. The optimization process directly affects the model's ability to discern relevant patterns and features from noisy data. Last but not least, regularization techniques serve as essential mechanisms to counter overfitting and promote robust generalization by, among other things, breaking the co-adaptation of neurons in the network. Understanding the synergy between normalization, optimization, and regularization and exploring the effect and impact of adaptive, hybrid methodologies on network performance, stability, and robustness are vital for training deep learning models capable of adapting to diverse data and achieving superior generalization performance. The insights presented in this research shed light on the interconnectedness of these key components, offering valuable guidance to enhance the generalization capabilities of deep learning and breaking the glass ceiling of current deep learning techniques. The research focuses on three key objectives:

* **AIM1 – Investigating Adaptive Strategies for Hybrid Data Normalization**: This objective aims to introduce an innovative adaptive approach to data normalization. To integrate additional knowledge into data-driven learning, a hybrid model for group normalization is proposed that dynamically regroups channels according to channel similarity. It can also handle groups of different sizes, in contrast to conventional group normalization. This strategy enhances efficiency and adaptability, and an extensive investigation will uncover the impact of this novel approach on model generalization.
* **AIM2 – Exploring High-Order EMA Optimizers:** Delving into optimization, this research will be the first to consider high-order Exponential Moving Average (EMA)-based optimizers. Our objective is to identify trends in network gradients better than conventional EMA-based optimizers, reducing lagging time and enhancing the stability and performance of deep learning models.
* **AIM3 – Studying Multi-level Dropout for Enhanced Regularization:** In this research, we will develop an adaptive dropout technique inspired by a combination of randomness, as characterized by the original dropout approach, and group dropout. Neurons will be grouped into clusters according to the spatial and semantic relationships among neighboring neurons, and then specific neurons will be randomly selected for dropout from each group during training. A multi-level regularization that consists of clustering and random processes will mitigate overfitting and reinforce model generalization.