Distinguishing between jets initiated by quarks and gluons is a challenging problem, yet very important for understanding the physics of the strong interaction. This thesis will present a novel Deep Learning approach to this problem using a neural network architecture based on the Transformer model trained on the jet constituents. We improve the existing architectures used in different tagging tasks and show their ability to accurately distinguish between quark and gluon jets. By combining techniques from different fields of Deep Learning, we propose a Dynamically Enhanced Particle Transformer (DeParT) that can surpass the state-of-the-art results in the quark/gluon jet tagging task.