Validation Loss Comparison of Different Model Dimensions (num_heads=4, dropout=0.1)