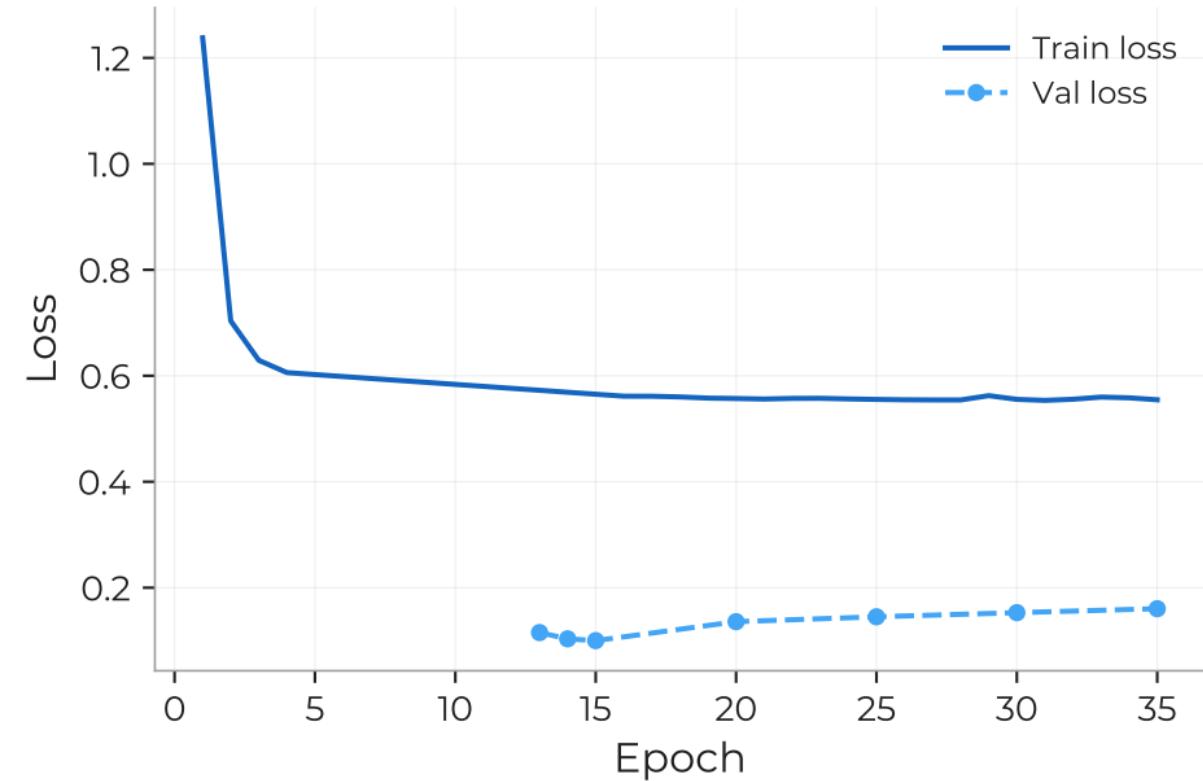


RetroTransformer v2 Training Curves

(a) Loss



(b) Token Accuracy

