Optimiser :ADAM with the AMSGRAD optimization method(Reddi, S.J., Kale, S., Kumar, S., 2018. On the convergence of ADAM

and beyond, in: International Conference on Learning Representations. pp. 1–23.)

The learning rate was initially set to 1e-4, and step decay policy was applied, specifically with [40, 30, 30, 20] epochs at the learning rate of [1e-4, 5e-5, 1e-5, 1e-6].







