 and  are estimates of the first moment (the mean) and the second moment (the uncentered variance) of the gradients respectively.

 and  are bias-corrected first and second moment estimates.

 and  are momentum coefficient.

 are gradients of the parameters we want to update.

are learning rate.

are smoothing term (avoids division by zero).

are weight.