parameters['W' + str(l+1)] -= learning\_rate\*grads['dW' + str(l+1)]

parameters['b' + str(l+1)] -= learning\_rate\*grads['db' + str(l+1)]

mini\_batch\_X = shuffled\_X[:, 0 : mini\_batch\_size]

mini\_batch\_Y = shuffled\_Y[:, 0 : mini\_batch\_size]

mini\_batch\_X = shuffled\_X[:, mini\_batch\_size \* num\_complete\_minibatches:]

mini\_batch\_Y = shuffled\_Y[:, mini\_batch\_size \* num\_complete\_minibatches:]

v["dW" + str(l+1)] = np.zeros(parameters["W" + str(l+1)].shape) #(numpy array of zeros with the same shape as parameters["W" + str(l+1)])

v["db" + str(l+1)] = np.zeros(parameters["b" + str(l+1)].shape) #(numpy array of zeros with the same shape as parameters["b" + str(l+1)])

v['dW' + str(l+1)] = beta\*v['dW' + str(l+1)] + (1-beta)\*grads['dW' + str(l+1)]

v['db' + str(l+1)] = beta\*v['db' + str(l+1)] + (1-beta)\*grads['db' + str(l+1)]

# update parameters

parameters['W' + str(l+1)] -= learning\_rate\*v['dW' + str(l+1)]

parameters['b' + str(l+1)] -= learning\_rate\*v['db' + str(l+1)]

v["dW" + str(l+1)] = np.zeros(parameters["W" + str(l+1)].shape) #(numpy array of zeros with the same shape as parameters["W" + str(l+1)])

v["db" + str(l+1)] = np.zeros(parameters["b" + str(l+1)].shape)#(numpy array of zeros with the same shape as parameters["b" + str(l+1)])

s["dW" + str(l+1)] = np.zeros(parameters["W" + str(l+1)].shape) #(numpy array of zeros with the same shape as parameters["W" + str(l+1)])

s["db" + str(l+1)] = np.zeros(parameters["b" + str(l+1)].shape) #(numpy array of zeros with the same shape as parameters["b" + str(l+1)])

v["dW" + str(l+1)] = beta1\*v["dW" + str(l+1)]+(1-beta1)\*grads['dW' + str(l+1)]

v["db" + str(l+1)] = beta1\*v["db" + str(l+1)]+(1-beta1)\*grads['db' + str(l+1)]

v\_corrected["dW" + str(l+1)] = v["dW" + str(l+1)]/(1 - (beta1\*\*t))

v\_corrected["db" + str(l+1)] = v["db" + str(l+1)]/(1 - (beta1\*\*t))

s["dW" + str(l+1)] = beta2\*s["dW" + str(l+1)] + (1-beta2)\*(grads['dW' + str(l+1)])\*\*2

s["db" + str(l+1)] = beta2\*s["db" + str(l+1)] + (1-beta2)\*(grads['db' + str(l+1)])\*\*2

s\_corrected["dW" + str(l+1)] = s["dW" + str(l+1)]/(1 - (beta2\*\*t))

s\_corrected["db" + str(l+1)] = s["db" + str(l+1)]/(1 - (beta2\*\*t))

parameters["W" + str(l+1)] -= (learning\_rate\*v\_corrected["dW" + str(l+1)])/(np.sqrt(s\_corrected["dW" + str(l+1)]+epsilon))

parameters["b" + str(l+1)] -= (learning\_rate\*v\_corrected["db" + str(l+1)])/(np.sqrt(s\_corrected["db" + str(l+1)]+epsilon))