

$$\mathbf{g} := \mathbb{E}_{\mathbf{x}, y \in B_i} \left[\frac{\partial \ell(\mathbf{x}, y | \theta)}{\partial \theta} \right]$$

$$\mathbf{m} := \alpha \mathbf{m} + (1 - \alpha) \mathbf{g}^2$$

$$\mathbf{v} := \rho \mathbf{v} + \frac{\mathbf{g}}{\sqrt{\mathbf{m} + \varepsilon}}$$

$$\theta := \theta - \varepsilon \mathbf{v}$$