

Exploiting Adam-like Optimization Algorithms to Improve the Performance of Convolutional Neural Networks

Loris Nanni¹, **Gianluca Maguolo**¹, Alessandra Lumini²

¹ University of Padova, Italy

² University of Bologna, Italy

Adam-like Ensemble

Adam algorithm

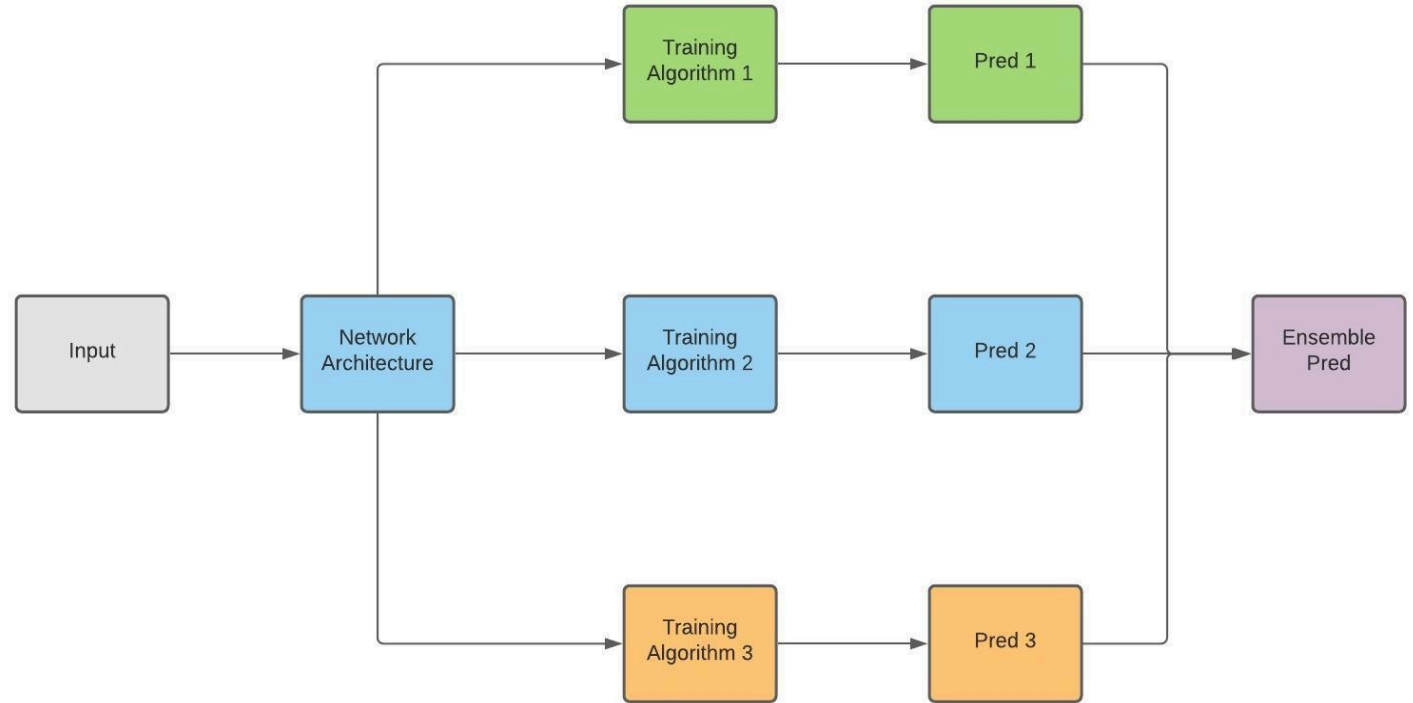
$$m_t = \rho_1 m_{t-1} + (1 - \rho_1) g_t$$

$$u_t = \rho_2 u_{t-1} + (1 - \rho_2) g_t^2$$

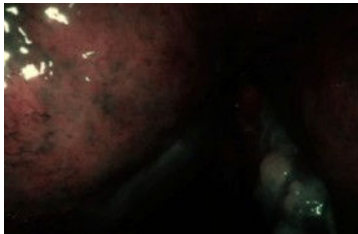
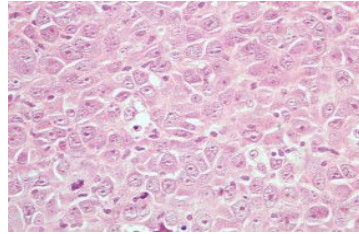
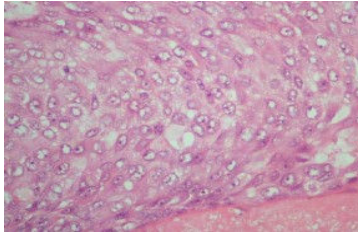
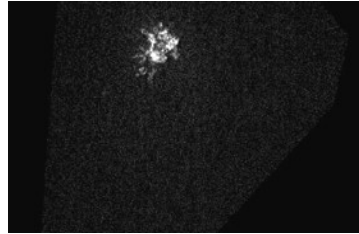
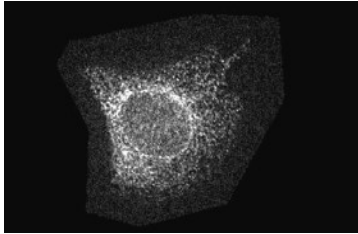
$$\hat{m}_t = \frac{m_t}{(1 - \rho_1^t)}$$

$$\hat{u}_t = \frac{u_t}{(1 - \rho_2^t)}$$

$$\theta_t = \theta_{t-1} - \lambda \frac{\hat{m}_t}{\sqrt{\hat{u}_t + \epsilon}}$$



Applications



On the left there are sample of the three datasets we used for testing (HeLa cells, Breast Carcinoma, Laryngeal Tissue).

Our ensemble established a new state of the art F1-score on the HeLa dataset.

Ensemble of multiple SGDs performs worse than a diverse ensemble