

How to Train Your Energy-Based Model for Regression

Fredrik K. Gustafsson¹ Martin Danelljan² Radu Timofte² Thomas B. Schön¹

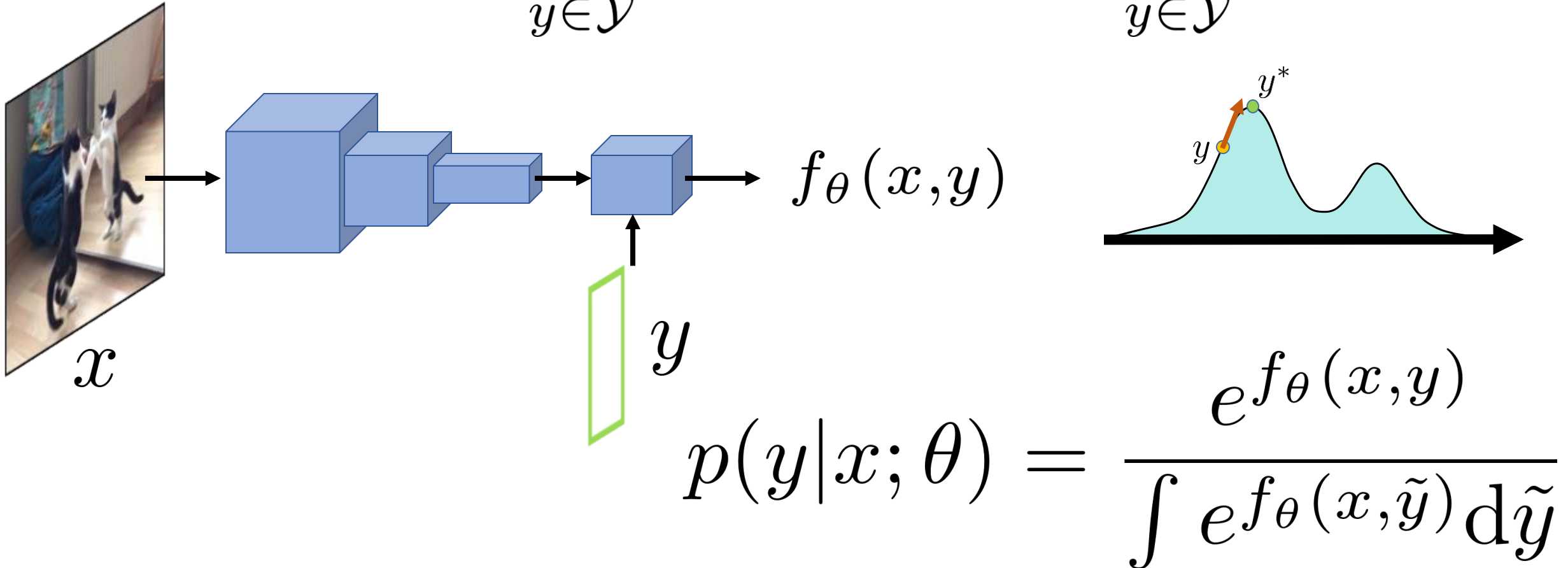
¹Department of Information Technology, Uppsala University, Sweden

²Computer Vision Lab, ETH Zürich, Switzerland



Energy-Based Models for Regression:

$$y^* = \arg \max_{y \in \mathcal{Y}} p(y|x; \theta) = \arg \max_{y \in \mathcal{Y}} f_{\theta}(x, y)$$

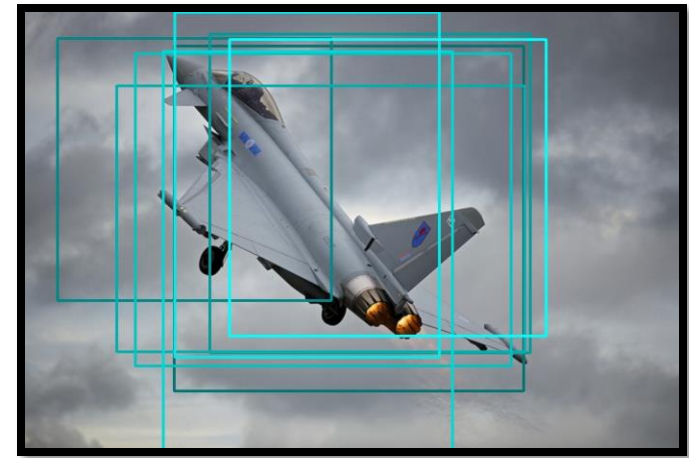


Noise Contrastive Estimation (NCE):

$$y^{(i,0)} \triangleq y_i$$



$$\{y^{(i,m)}\}_{m=1}^M \sim p_N(y|y_i)$$



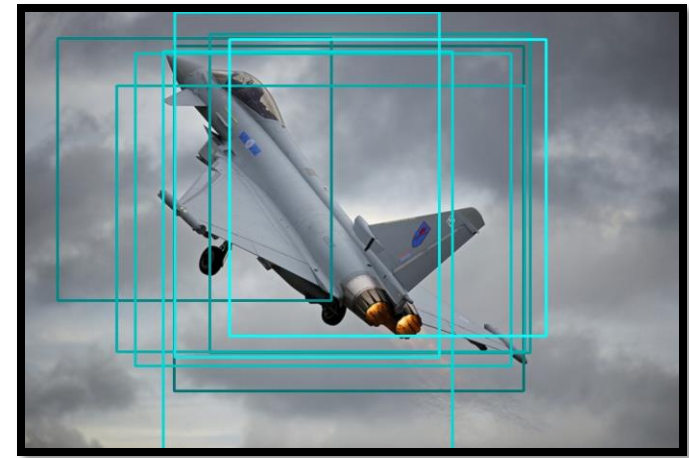
$$J(\theta) = -\frac{1}{n} \sum_{i=1}^n \log \frac{\exp \{ f_{\theta}(x_i, y^{(i,0)}) - \log p_N(y^{(i,0)} | y_i) \}}{\sum_{m=0}^M \exp \{ f_{\theta}(x_i, y^{(i,m)}) - \log p_N(y^{(i,m)} | y_i) \}}$$

Our Proposed Training Method (NCE+):

$$y^{(i,0)} \triangleq y_i + \nu_i$$



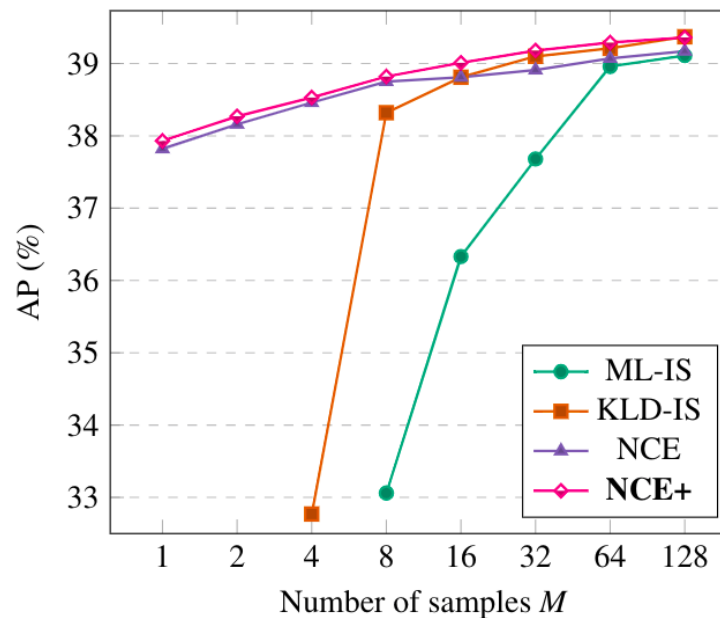
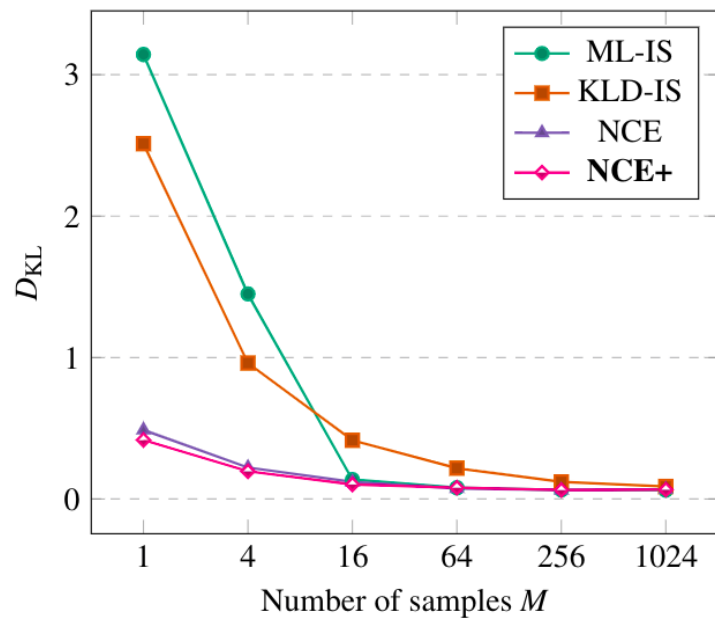
$$\{y^{(i,m)}\}_{m=1}^M \sim p_N(y|y_i)$$



$$J(\theta) = -\frac{1}{n} \sum_{i=1}^n \log \frac{\exp \{ f_{\theta}(x_i, y^{(i,0)}) - \log p_N(y^{(i,0)} | y_i) \}}{\sum_{m=0}^M \exp \{ f_{\theta}(x_i, y^{(i,m)}) - \log p_N(y^{(i,m)} | y_i) \}}$$

Comparison of Training Methods:

	ML-IS	ML-MCMC-1	ML-MCMC-4	ML-MCMC-8	KLD-IS	NCE	DSM	NCE+
AP (%) \uparrow	39.4	36.4	36.4	36.4	39.6	39.5	36.3	39.7
AP ₅₀ (%) \uparrow	58.6	57.9	57.9	58.0	58.6	58.6	57.9	58.7
AP ₇₅ (%) \uparrow	42.1	38.8	39.0	39.0	42.6	42.4	38.9	42.7
Training Cost \downarrow	1.03	2.47	7.05	13.3	1.02	1.04	3.84	1.09



Visual Tracking Experiments:

	MDNet [40]	UPDT [4]	DaSiamRPN [60]	ATOM [8]	SiamRPN++ [30]	DiMP [5]	SiamRCNN [53]	PrDiMP [9]	DiMP- KLD-IS	DiMP- NCE	DiMP- NCE+
TrackingNet	60.6	61.1	63.8	70.3	73.3	74.0	81.2	75.8	78.1	77.1	78.7
LaSOT	39.7	-	-	51.5	49.6	56.9	64.8 (62.3)	59.8	63.1	62.8	63.7
UAV123	52.8	54.5	57.7	63.2	61.3	64.3	64.9	66.7	66.6	65.2	67.2
NFS	42.2	53.7	-	58.4	-	62.0	63.9	63.5	64.7	64.3	65.0
OTB-100	67.8	70.2	65.8	66.9	69.6	68.4	70.1 (68.0)	69.6	70.1	69.3	70.7



github.com/fregu856/ebms_regression

github.com/visionml/pytracking