

Supplementary for CERN: Confidence-Energy Recurrent Network for Group Activity Recognition

Tianmin Shu¹, Sinisa Todorovic² and Song-Chun Zhu¹

¹ University of California, Los Angeles ²Oregon State University

tianmin.shu@ucla.edu sinisa@onid.orst.edu sczhu@stat.ucla.edu

6. The Energy Layer of CERN

The regularized energy $\tilde{\mathcal{E}}(X, Y, c)$ can be reformulated in a compact form as

$$\begin{aligned} \tilde{\mathcal{E}}(X, Y, c) = & \mathbf{w}_c^V \top \boldsymbol{\psi}^V - \boldsymbol{\lambda}^V \top \log \mathbf{p}_c^V \\ & + \mathbf{w}_c^E \top \boldsymbol{\psi}^E - \boldsymbol{\lambda}^E \top \log \mathbf{p}_c^E \\ & + w_c \psi_c - \lambda \log p_c, \end{aligned} \quad (1)$$

where weights of the EL are grouped into $\{w_c\}_{c=1, \dots, C}$, λ , and the following parameter vectors:

$$\begin{aligned} \mathbf{w}_c^V = & [w_{c,1}^V, \dots, w_{c,|\mathcal{Y}^V|}^V]^\top, \mathbf{w}_c^E = [w_{c,1}^E, \dots, w_{c,|\mathcal{Y}^E|}^E]^\top, \\ \boldsymbol{\lambda}^V = & [\lambda^V, \dots, \lambda^V]^\top, \quad \boldsymbol{\lambda}^E = [\lambda^E, \dots, \lambda^E]^\top, \end{aligned} \quad (2)$$

and the input to the EL is specified in terms of the LSTM softmax outputs and p-values:

$$\begin{aligned} \boldsymbol{\psi}^V = & \left[\sum_{i:y_i=1} \psi^V(x_i, y_i), \dots, \sum_{i:y_i=|\mathcal{Y}^V|} \psi^V(x_i, y_i) \right]^\top, \\ \boldsymbol{\psi}^E = & \left[\sum_{\substack{(i,j): \\ y_{ij}=1}} \psi^E(x_{ij}, y_{ij}), \dots, \sum_{\substack{(i,j): \\ y_{ij}=|\mathcal{Y}^E|}} \psi^E(x_{ij}, y_{ij}) \right]^\top, \\ \psi_c = & \psi(x, c), \\ \mathbf{p}_c^V = & \left[\sum_{i:y_i=1} p_i^V(c, y_i), \dots, \sum_{i:y_i=|\mathcal{Y}^V|} p_i^V(c, y_i) \right]^\top, \\ \mathbf{p}_c^E = & \left[\sum_{(i,j):y_{ij}=1} p_{ij}^V(c, y_i), \dots, \sum_{(i,j):y_{ij}=|\mathcal{Y}^E|} p_{ij}^V(c, y_i) \right]^\top, \\ p_c = & p(c). \end{aligned} \quad (3)$$