Method Overview

This paper proposes a knowledge-guided fashion network to solve the problem of visual fashion analysis, e.g., fashion landmark localization and clothing category classification.

Contribution

i) A deep grammar network to encode a set of knowledge over fashion clothes.

ii) A Bidirectional Convolutional Recurrent Neural Network (BCRNN) for approaching message passing over the fashion grammars.

iii) Two fashion attention mechanisms, one is landmark-aware and domain-knowledge-involved, and the other one directly focuses on the category relevant image regions and can be learned in goal driven manner.

Proposed Fashion Model

(a) Input fashion image. (b) A set of BCRNNs (yellow cubes) are established for capturing kinematics and symmetry grammars as global constraints for detecting clothing landmarks (blue cubes). Fashion landmark-aware attention and clothing category-driven attention (red cubes are further incorporated for enhancing clothing features and improving clothing category classification and attribute estimation).

(c) Results for clothing landmark detection, category classification and attribute estimation.

EXPERIMENTS

Performance for Fashion Category Classification and Attribute Prediction over DeepFashion-C dataset

Performance for clothing landmark detection over DeepFashion-C (top) and FLD (bottom) datasets

Clothing category classification results (top) and visualization of attention mechanisms (bottom) on DeepFashion-C dataset

Fashion Landmark-Aware Attention $A_L$, obtained via applying cross-channel average-pooling operation over the landmark heatmaps.

Clothing Category-Driven Attention $A_C$, learned as an implicit, top-down network attention mechanism.

Let $F$ denote the fashion feature from our fashion network, it can be enhanced via fashion attentions: $G = (1+A_L^c+A_C^c)*F$