SizeNet: Weakly Supervised Learning of Visual Size and Fit in Fashion Images

Nour Karessli Romain Guigourès Reza Shirvany

Zalando SE - Berlin, Germany

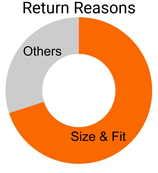
{nour.karessli, romain.guigoures, reza.shirvany}@zalando.de

CVPR LONG BEACH CALIFORNIA June 16-20, 2019

Motivation

Finding clothes that fit is the biggest problem for customers shopping fashion online and offline. Supporting customers on their size and fit purchase decision in e-commerce context is particularly challenging:

- Thousands of new articles get activated everyday with short lifetime
- Return process takes from few days to few weeks resulting in zero or a few sales and returns data points for new articles



Contributions

In this paper we introduce a novel teacher-student approach on fashion images to:

- Investigate and demonstrates the rich value of fashion images in inferring size characteristics of fashion apparel
- Effectively tackle the challenging cold start problem of providing size advice for new articles with zero/few return data
- Generate large scale confidence-weighted weak annotations from crowd's subjective feedback- enabling us to control the influence of weak annotations on the final model

Related Work

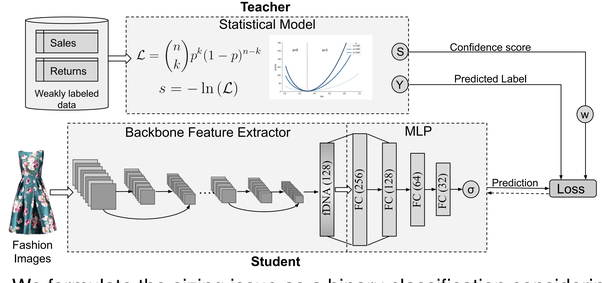
Teacher-Student Transfer Learning

Transferring knowledge from privileged information space to decision space [1]:

- Teacher leverages privileged historical weakly annotated data of *sales and returns*
- Student uses this knowledge to learn from *images* in decision space

[1] V Vapnik and R Izmailov. Learning using privileged information: Similarity control and knowledge transfer. In *JMLR15*.
[2] K He, X Zhang, S Ren, and J Sun. Deep residual learning for image recognition. In *CVPR16*.
[3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *CVPR09*.
[4] C Bracher, S Heinz, and R Vollgraf. Fashion dna: Merging content and sales data for recommendation and article mapping. In *KDD16*.
[5] V Petsiuk, A Das, and K Saenko. Rise: Randomized input sampling for explanation of black-box models. In *BMVC18*.

SizeNet: Learning Visual Size Cues



We formulate the sizing issue as a binary classification considering that:

- Article categories have different size related return rate
- Article sales period influences the return rate

TEACHER

Statistical Model: $\mathcal{L} = \binom{n}{k} p^k (1-p)^{n-k}$.

$\mathcal{L}$: binomial likelihood, p: expected size return rate of article category, k: number of article size returns, n: number of article sales. The estimator score is define as $s = -\ln(\mathcal{L})$

CNN Backbone Feature Extractor

Transfer knowledge using bottleneck features of pre-trained network

- ResNet [2] pre-trained on ImageNet [3] dataset

STUDENT

- FashionDNA [4] Resnet-like pre-trained on rich in-house fashion dataset

MLP: Multi-Layer Perceptron

4 fully connected layers with nonlinear activations. Use weighted binary cross entropy loss based on estimator confidence score $w = \ln(1+s)$; the logarithmic transformation reduces the score skewness

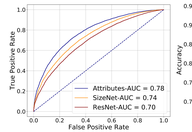
Evaluation

Dataset Women textile including 12 categories such as dresses, blouses, jeans, skirts, etc.

Class	#Articles	# Images
size issue	68,892	69,064
no size issue	58,152	58,321
total	127,044	127,385

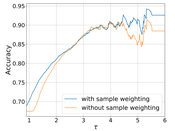
Attributes Baseline

Replace article images with sparse k-hot encoding of *human annotated* binary fashion attributes → fashion images achieve comparable results



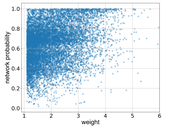
Weights Importance

Exploiting weights offers better generalization capacity



Student Prediction vs. Teacher Confidence

BOTTOM RIGHT almost no samples are misclassified by SizeNet when Teacher is certain of no size issue; **TOP LEFT** high density of correctly predicted samples by SizeNet where Teacher is unsure; **TOP RIGHT** samples show that SizeNet has learned accurately from Teacher; **BOTTOM LEFT** SizeNet misclassifies fewer samples where Teacher is unsure



Size Issue Explanations using RISE [5]

True Positives show localized heatmaps where False Positives are affected by article design