Category-Specific CNN for Visual-aware CTR Prediction at JD.com

KDD '20, August 23–27, 2020,

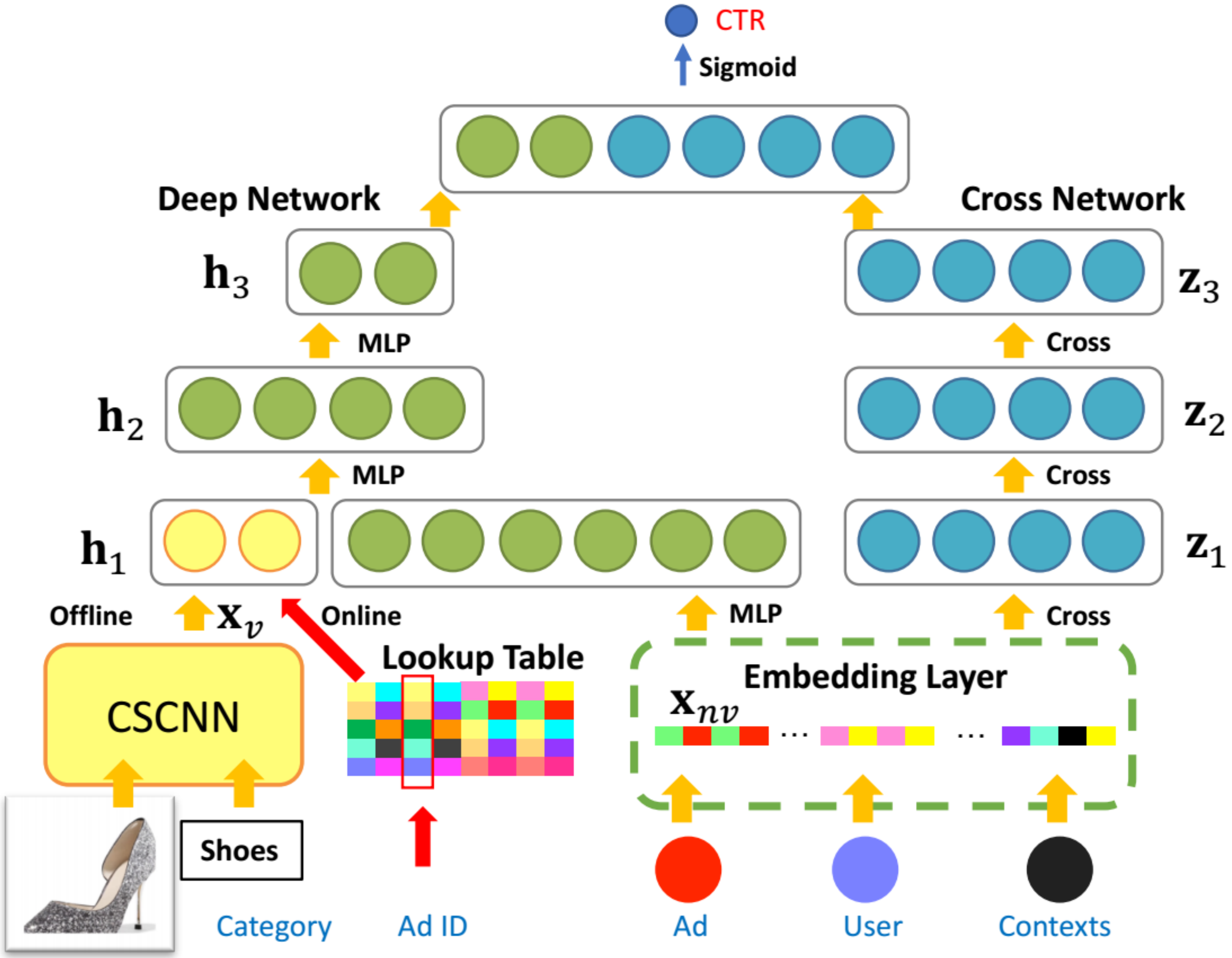
problem

- the strict requirements for efficient end-to-end training and low-latency online serving.

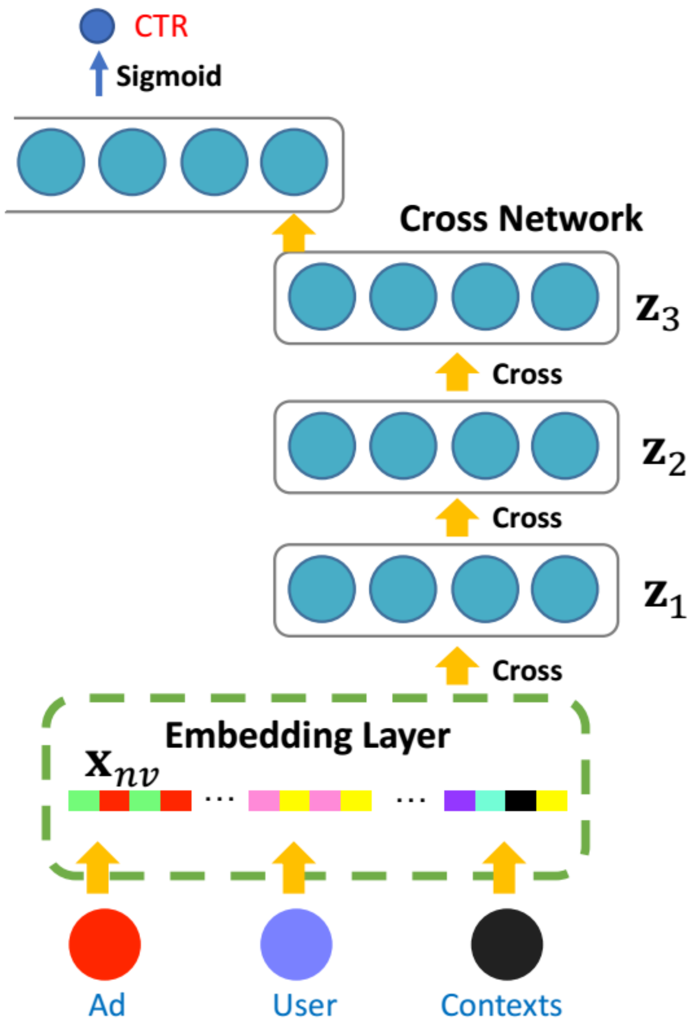
- the off-the-shelf CNNs and late fusion architectures are suboptimal.

- off-the-shelf CNNs were designed for classification thus never take categories as input features.

CTR Prediction System



Cross Network



The cross net is used to process non-visual feature

$$z_{l+1} = z_0 z_l^T w_l + b_l + z_l$$

$$\mathrm{z}_0 = \mathrm{x}_{\mathrm{nv}}$$

$$\mathrm{X}_{\mathrm{nv}} = \mathbb{R}^{380}$$

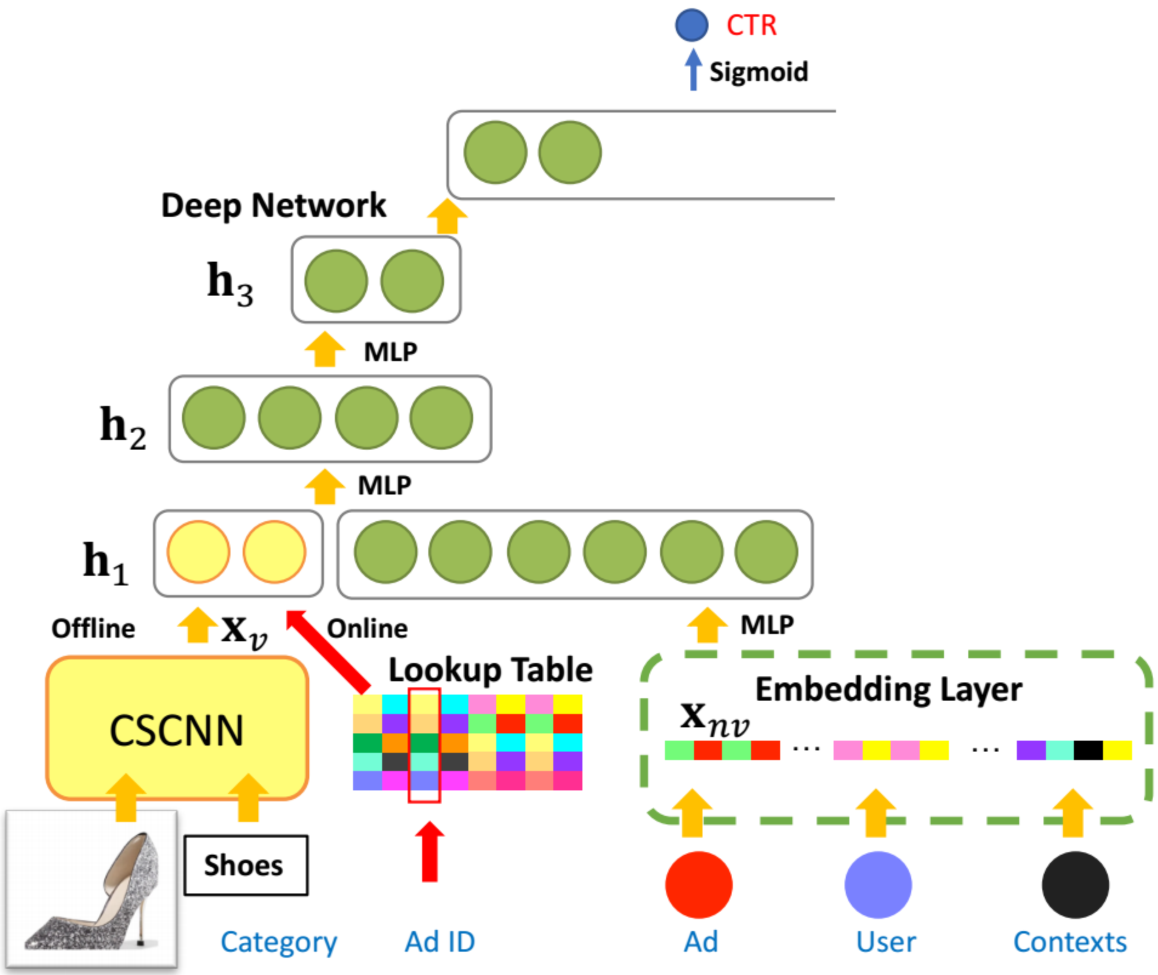
Why is 380?

$$x_{emb} = E x_{hot}$$

where $E \in \mathbb{R}^{d_e \times v}$ is the embedding dictionary for this specific feature and d_e is the embedding size. We then concatenate the x_{emb}'s of all features to build x_{nv}.

Setting $d_e = 4$, the total dimension is $95 \times 4 = 380$.

Deep Network



In layer 1, we transform the non-visual feature to 1024 dimension and concatenate it with the visual feature,

$$\mathrm{h}_1 = [x_v, ReLU - MLP(x_{nv})] \in \mathbb{R}^{150+1024}$$

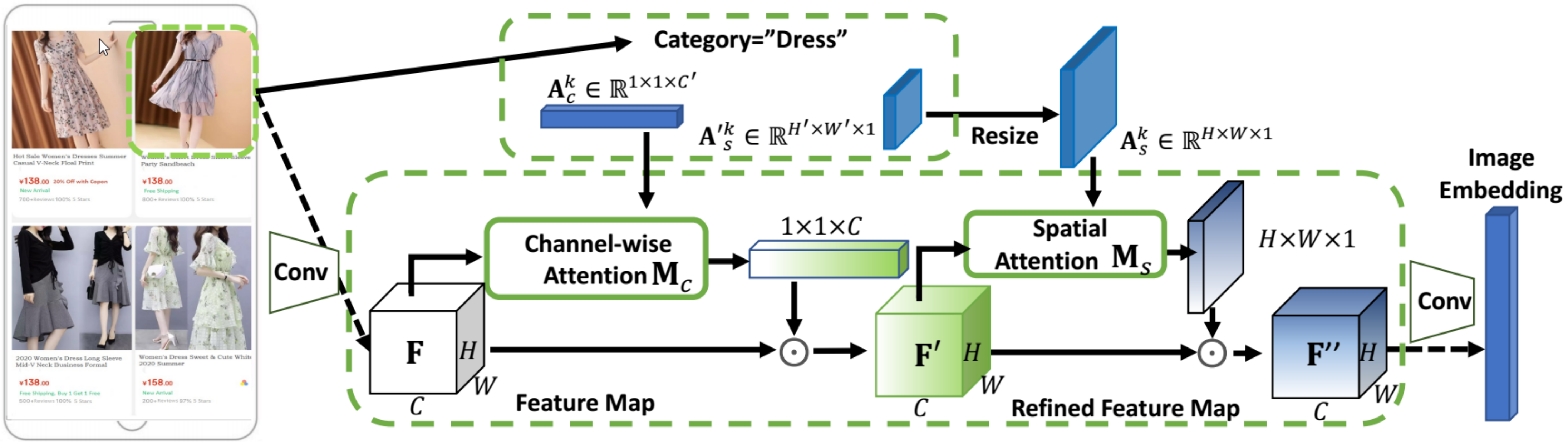
Two deep layers follows,

$$h_{l+1} = ReLU - MLP(h_l), l \in \{1,2\}, h_2 \in \mathbb{R}^{512}, h_3 \in \mathbb{R}^{256}$$

Finally, we combine the outputs for the predicted CTR,

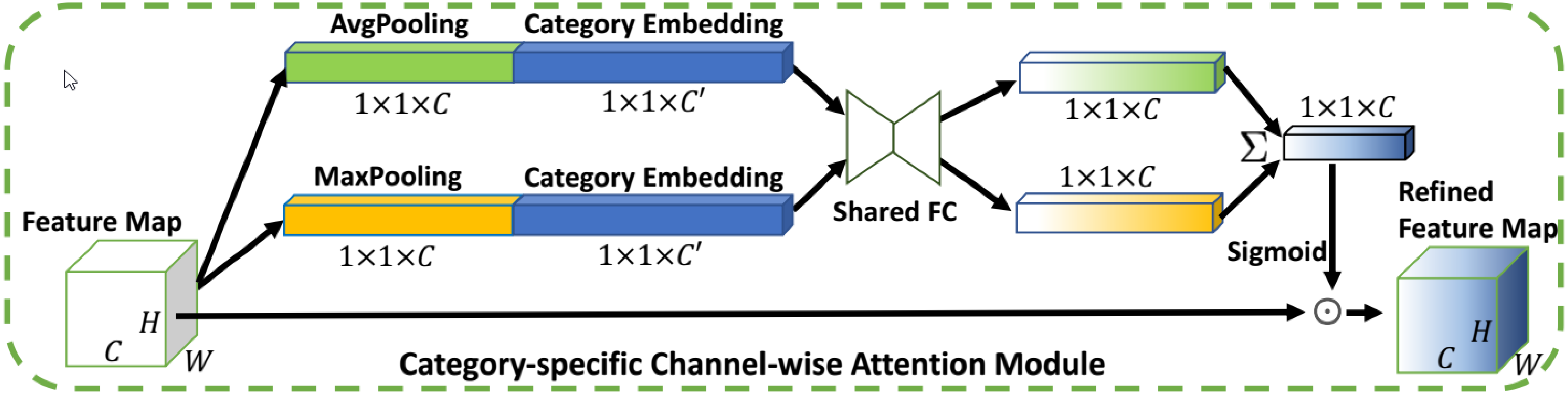
$$\hat{y} = \sigma(ReLU - MLP[h_3, z_3])$$

Category-Specific CNN(CSCNN)



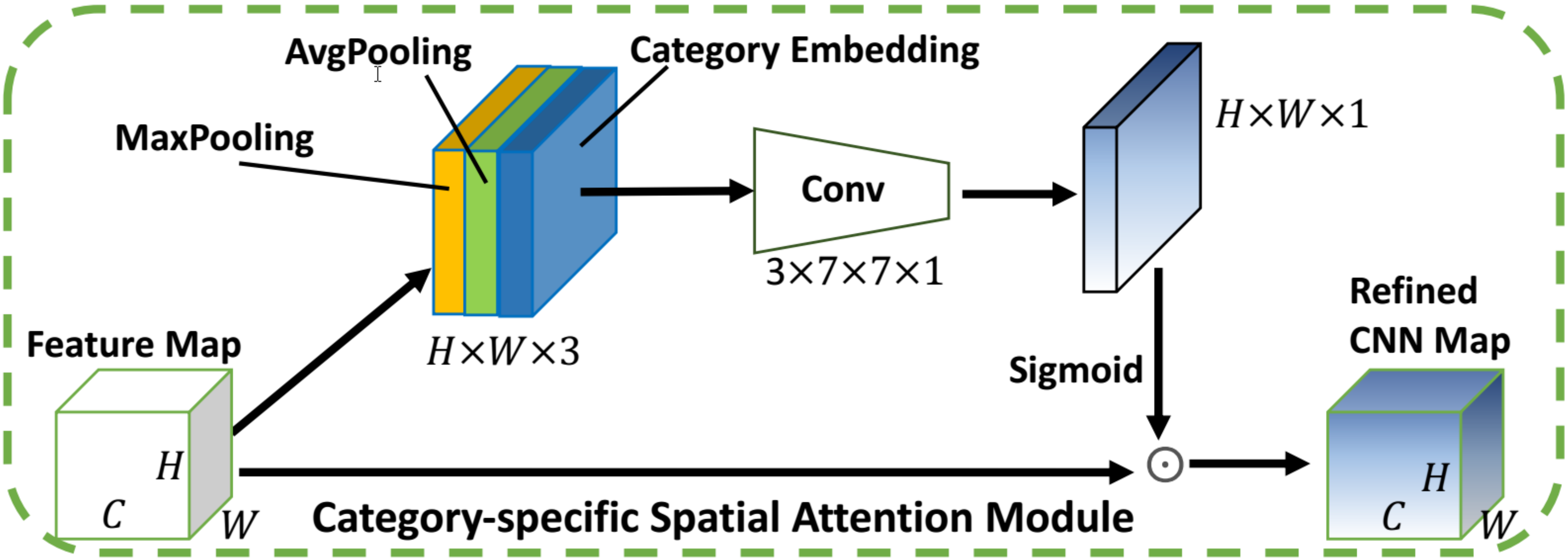
$$F' = M_c(F, A_c^k) \odot F \qquad F'' = M_s(F', A_s^k) \odot F'$$

Category-specific Channel-wise Attention



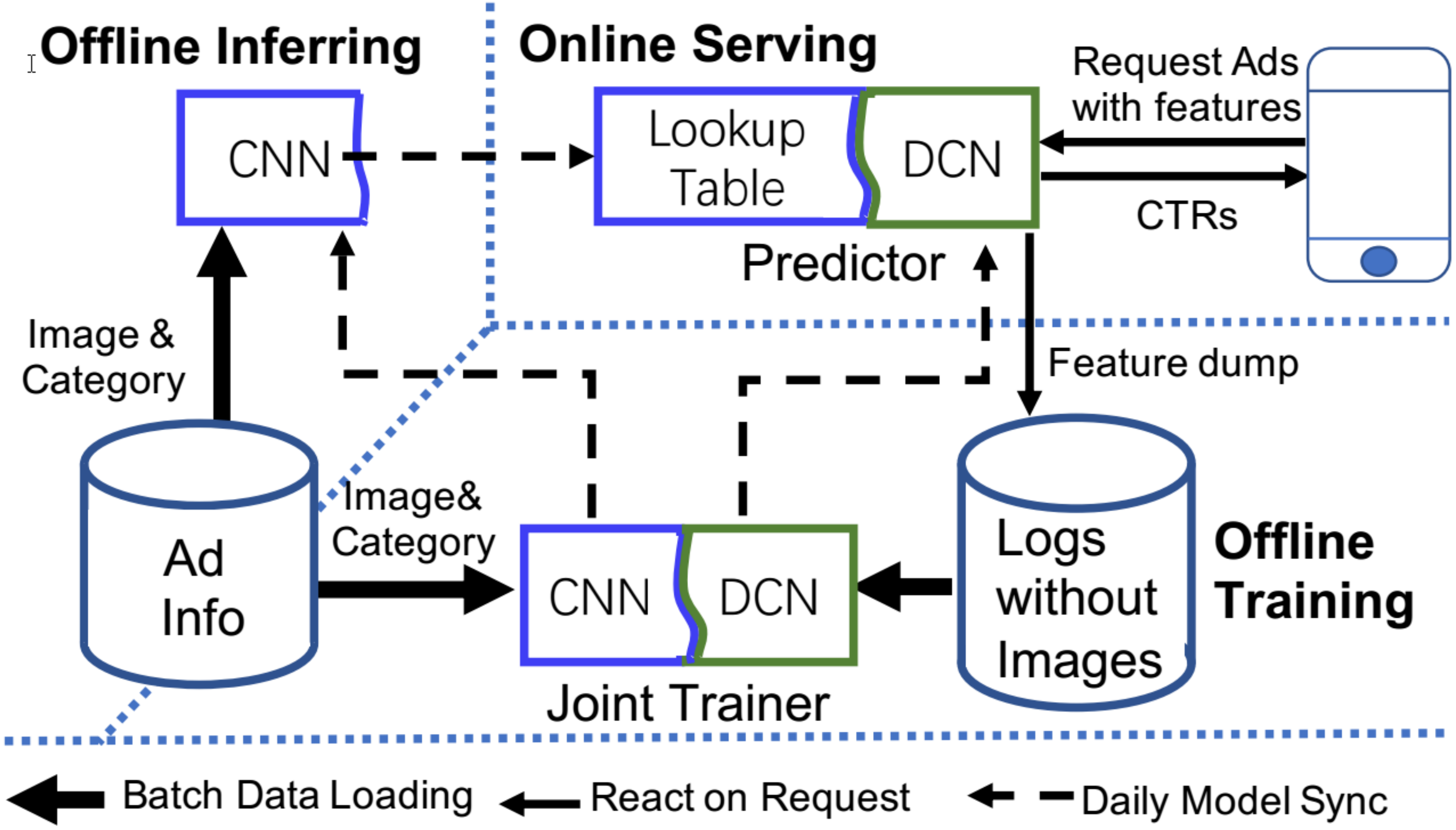
$$M_c(F, A_c^k) = \sigma\left(MLP[AvgP(F), A_c^k] + MLP[MaxP(F), A_c^k]\right),$$

Category-specific Spatial Attention



$$M_s(F', A_s^k) = \sigma(Conv_{7\times7}(MaxP(F'), AvgP(F'), A_s^k))$$

Online model system



Datasets

- The ablation study is conducted on 3 wildly used benchmark datasets about products on Amazon.com introduced in [16].

New: Amazon 2018 dataset

We've put together a new version of our Amazon data, including more reviews and additional metadata

New: Advice to Prospective Students

If you are considering internships, PhD applications, or project work, please read this advice first before contacting me about joining my lab

New: Advice to Students Requesting Reference Letters

New: Repository of Recommender Systems Datasets

A collection of datasets for recommender systems research is now available on our lab's dataset webpage

Datasets

Dataset	#Users	#Items	# Interact	#Category
Fashion	64,583	234,892	513,367	49
Women	97,678	347,591	827,678	87
Men	34,244	110,636	254,870	62

Evaluation Metrics

- AUC measures the probability that a randomly sampled positive item has higher preference than a sampled negative one,

$$AUC = \frac{1}{|\mathcal{U}|}\sum_{u\in\mathcal{U}}\frac{1}{|\mathcal{D}_u|}\sum_{(i,j)\in\mathcal{D}_u}\mathbb{I}(\hat{y}_{u,i} > \hat{y}_{u,j})$$

$$\mathcal{D}_{\mathrm{u}} = \{(i,j)|(u,i) \in \mathcal{T}_u \; and \; (u,j) \notin (\mathcal{P}_u \cup \mathcal{V}_u \cup \mathcal{T}_u)\}$$

- In JD.com, every 1‰increase in off-line AUC brings 6 million dollars lift in the overall advertising income.

Experimental

Datasets		No Image	With Image		With Image + Category				
		BPR-MF	VBPR	DVBPR	DVBPR-C	Sherlock	DeepStyle	DVBPR-SCA	Ours
Fashion	All	0.6147	0.7557	0.8011	0.8022	0.7640	0.7530	0.8032	**0.8156**
	Cold	0.5334	0.7476	0.7712	0.7703	0.7427	0.7465	0.7694	**0.7882**
Women	All	0.6506	0.7238	0.7624	0.7645	0.7265	0.7232	0.7772	**0.7931**
	Cold	0.5198	0.7086	0.7078	0.7099	0.6945	0.7120	0.7273	**0.7523**
Men	All	0.6321	0.7079	0.7491	0.7549	0.7239	0.7279	0.7547	**0.7749**
	Cold	0.5331	0.6880	0.6985	0.7018	0.6910	0.7210	0.7048	**0.7315**

When testing, we report performance on two sets of items: All items, and Cold items with fewer than 5 actions in the training set.

Experimental

	Original		+CSCNN	
	All	Cold	All	Cold
No Attention	0.7491	0.6985	–	–
SE	0.7500	0.6989	**0.7673**	**0.7153**
CBAM-Channel	0.7506	0.7002	**0.7683**	**0.7184**
CBAM-All	0.7556	0.7075	**0.7749**	**0.7315**

Experimental

		CNN-F	Inception
No Attention	All	0.7491	0.7747
	Cold	0.6985	0.7259
CBAM	All	0.7556	0.7794
	Cold	0.7075	0.7267
CSCNN	All	**0.7749**	**0.7852**
	Cold	**0.7315**	**0.7386**

What is new

- We proposed Category-specific CNN, specially designed for visual-aware CTR prediction in e-commerce.
- Our early-fusion architecture enables category-specific feature recalibration and emphasizes features that are both important and category related, which contributes to significant performance gain in CTR prediction tasks.
- With the help of a highly efficient infrastructure, CSCNN has now been deployed in the search advertising system of JD.com, serving the main traffic of hundreds of millions of active users.

Thank you