Image Retrieval via Canonical Correlation Analysis

Kangdi Shi[†], Xiaohong Liu[†], Muhammad Alrabeiah[†], Xingtong Guo[†],
Jie Lin[‡], Huan Liu[†], Jun Chen[†]

[†] McMaster University, Ontario, Canada

[‡] Institute of Information Research, A-STAR, Singapore

16th Canadian Workshop on Information Theory, 2019

ENGINEERING McMaster University

Image Retrieval via Canonical Correlation Analysis

- The basic image retrieval process based on convolutional neutral network.

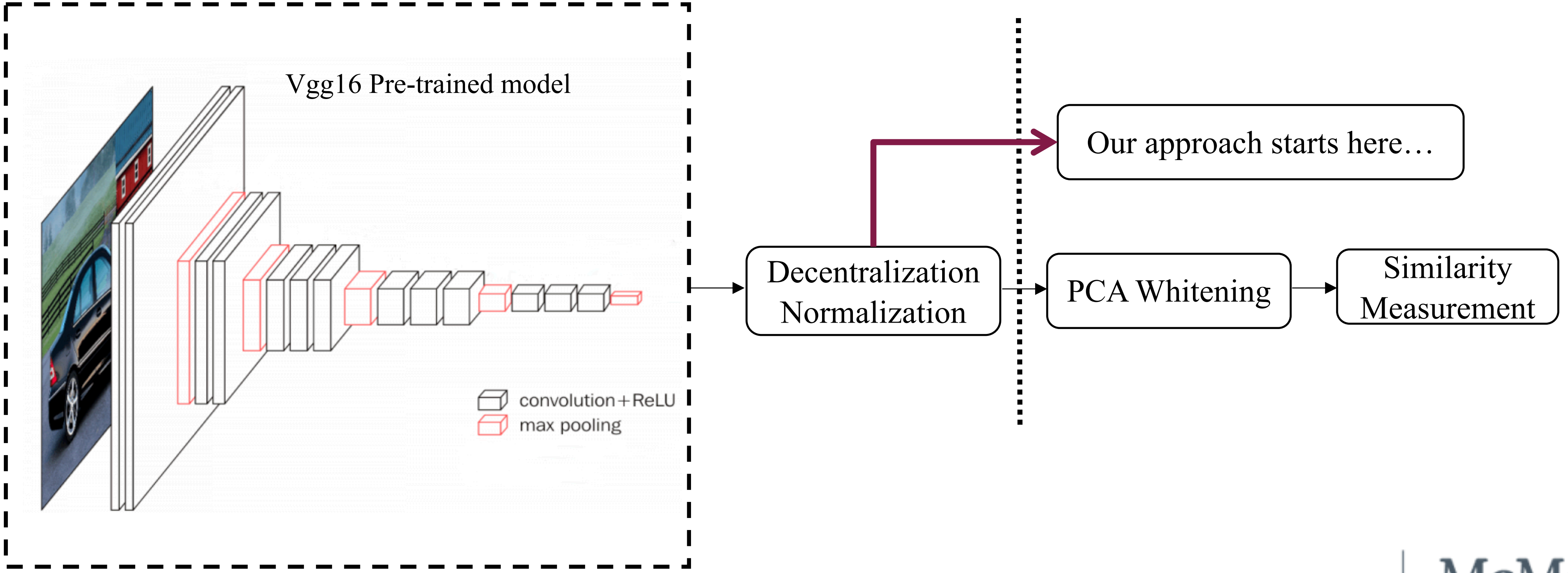

Image Retrieval via Canonical Correlation Analysis

- Dataset Structure

  - Matching pairs:
    - $\mathbf{X}^{M} = [\, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 , \mathbf{x}_4 \ldots, \mathbf{x}_L]_{512xL}$
    - $\mathbf{Y}^{M} = [\, \mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4 \ldots, \mathbf{y}_L]_{512xL}$

  - Non-matching pairs:
    - $\mathbf{X}^{N} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 , \mathbf{x}_4 \ldots, \mathbf{x}_L]_{512xL}$
    - $\mathbf{Y}^{N} = [\mathbf{y}_4, \mathbf{y}_8, \mathbf{y}_1, \mathbf{y}_{10} \ldots, \mathbf{y}_{L-4}]_{512xL}$

  - $\mathbf{x}_i$ matches $\mathbf{y}_j$ when i = j. Otherwise, they doesn't match.

Image Retrieval via Canonical Correlation Analysis

- Correlation analysis and canonical vectors
  - Step 1 – Building covariance matrix :

$$\mathbf{\Phi}^M = \frac{1}{2(L-1)} \mathbf{H}^M (\mathbf{H}^M)^T = \begin{bmatrix} \Sigma_{auto} & \Sigma^M \\ \Sigma^M & \Sigma_{auto} \end{bmatrix}, \text{ where } \mathbf{H}^M = \begin{bmatrix} X^M & Y^M \\ Y^M & X^M \end{bmatrix}_{1024\text{x}2\text{L}}$$

$$\mathbf{\Phi}^N = \frac{1}{2(L-1)} \mathbf{H}^N (\mathbf{H}^N)^T = \begin{bmatrix} \Sigma_{auto} & \Sigma^N \\ \Sigma^N & \Sigma_{auto} \end{bmatrix}, \text{ where } \mathbf{H}^N = \begin{bmatrix} X^N & Y^N \\ Y^N & X^N \end{bmatrix}_{1024\text{x}2\text{L}}$$

Note: $\Sigma_{auto} = \frac{X^M (X^M)^T + Y^M (Y^M)^T}{2(L-1)} = \frac{X^N (X^N)^T + Y^N (Y^N)^T}{2(L-1)}$,

$$\Sigma^M = \frac{X^M (Y^M)^T + Y^M (X^M)^T}{2(L-1)}, \; \Sigma^N = \frac{X^N (Y^N)^T + Y^N (X^N)^T}{2(L-1)}.$$

Image Retrieval via Canonical Correlation Analysis

- Correlation analysis and canonical vectors
  - Step 2 – Removing autocorrelation:

$$\widehat{\boldsymbol{\Phi}}^{M} = \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}} \boldsymbol{\Phi}^{M} \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}} = \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}} \begin{bmatrix} \boldsymbol{\Sigma}_{auto} & \boldsymbol{\Sigma}^{M} \\ \boldsymbol{\Sigma}^{M} & \boldsymbol{\Sigma}_{auto} \end{bmatrix} \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}} = \begin{bmatrix} \mathbf{I} & \mathbf{J}^{M} \\ \mathbf{J}^{M} & \mathbf{I} \end{bmatrix},$$

$$\widehat{\boldsymbol{\Phi}}^{N} = \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}} \boldsymbol{\Phi}^{N} \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}} = \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}} \begin{bmatrix} \boldsymbol{\Sigma}_{auto} & \boldsymbol{\Sigma}^{N} \\ \boldsymbol{\Sigma}^{N} & \boldsymbol{\Sigma}_{auto} \end{bmatrix} \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}} = \begin{bmatrix} \mathbf{I} & \mathbf{J}^{N} \\ \mathbf{J}^{N} & \mathbf{I} \end{bmatrix}.$$

Where $\mathbf{J}^{M} = \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}} \boldsymbol{\Sigma}^{M} \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}}$, $\mathbf{J}^{N} = \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}} \boldsymbol{\Sigma}^{N} \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}}$.

Image Retrieval via Canonical Correlation Analysis

- Correlation analysis and canonical vectors
  - Step 3 – Deriving matching and non-matching coefficients :

    1) Applying eigen-decomposition on $\mathbf{J}^M$:

$$\mathbf{J}^M = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\boldsymbol{T}}.$$

    2) Left- and right- multiplication of $\mathbf{J}^M$ and $\mathbf{J}^N$ by $\mathbf{U}^T$ and $\mathbf{U}$:

$$\boldsymbol{\Lambda} = \mathbf{U}^T\mathbf{J}^M\mathbf{U};\ \boldsymbol{\Pi} = \mathbf{U}^T\mathbf{J}^N\mathbf{U}.$$

Image Retrieval via Canonical Correlation Analysis

- Correlation analysis and canonical vectors
  - Step 3 – Deriving matching and non-matching coefficients :

$$\boldsymbol{\Lambda} = \begin{bmatrix} c_1^M & 0 & \cdots & 0 \\ 0 & c_1^M & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_{512}^M \end{bmatrix}, \boldsymbol{\Pi} = \begin{bmatrix} c_1^N & \pi_{1,2} & \cdots & \pi_{1,512} \\ \pi_{1,2} & c_2^N & \cdots & \pi_{2,512} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{512,1} & \pi_{512,1} & \cdots & c_{512}^N \end{bmatrix}$$

Matching coefficient: $c_t^M = [\boldsymbol{\Lambda}]_{tt}$,
Non-matching coefficient: $c_t^N = [\boldsymbol{\Pi}]_{tt}$, where $t \in \{1,2,\ldots,512\}$.

Image Retrieval via Canonical Correlation Analysis

- Chernoff information for canonical vector selection

Step 1 – The Chernoff information $\mathrm{CI}(\boldsymbol{S}_t^M \| \boldsymbol{S}_t^N)$ is defined as

$$\mathrm{CI}(\boldsymbol{S}_t^M \| \boldsymbol{S}_t^N) = \mathrm{D}(\boldsymbol{S}_t^\lambda \parallel \boldsymbol{S}_t^M) = \mathrm{D}(\boldsymbol{S}_t^\lambda \parallel \boldsymbol{S}_t^M),$$

where $\boldsymbol{S}_t^M = \begin{bmatrix} 1 & c_t^M \\ c_t^M & 1 \end{bmatrix}, \boldsymbol{S}_t^N = \begin{bmatrix} 1 & c_t^M \\ c_t^M & 1 \end{bmatrix},$

$$(S_t^\lambda)^{-1} = \lambda_t (\boldsymbol{S}_t^M)^{-1} + (1 - \lambda_t)\,(\boldsymbol{S}_t^N)^{-1},$$

$\lambda_t \in [0,1]$, t $\in [1,512]$.

ENGINEERING McMaster University

Image Retrieval via Canonical Correlation Analysis

- Chernoff information for canonical vector selection

Step 2 – We express the "distance" between distributions $\boldsymbol{S}_t^{\lambda}$ and $\boldsymbol{S}_t^{M}$, $\boldsymbol{S}_t^{N}$ by using

Kullback–Leibler (KL) divergence:

$$\mathrm{D}(\boldsymbol{S}_t^{\lambda} \,||\, \boldsymbol{S}_t^{M}) = \frac{1}{2}\log_e \frac{|\boldsymbol{S}_t^{M}|}{|S_t^{\lambda}|} + \frac{1}{2}\,\mathrm{tr}(\,(\boldsymbol{S}_t^{M})^{-1} S_t^{\lambda}) \text{ -1},$$

$$\mathrm{D}(\boldsymbol{S}_t^{\lambda} \,||\, \boldsymbol{S}_t^{N}) = \frac{1}{2}\log_e \frac{|\boldsymbol{S}_t^{N}|}{|S_t^{\lambda}|} + \frac{1}{2}\,\mathrm{tr}(\,(\boldsymbol{S}_t^{N})^{-1} S_t^{\lambda}) \text{ -1}.$$

Setting $\mathrm{D}(\boldsymbol{S}_t^{\lambda} \,||\, \boldsymbol{S}_t^{M}) = \mathrm{D}(\boldsymbol{S}_t^{\lambda} \,||\, \boldsymbol{S}_t^{N})$ helps find the optimal $\lambda_t = \lambda_t^*$.

Image Retrieval via Canonical Correlation Analysis

- Chernoff information for canonical vector selection
  - Step 3 – Feature selection:

    By solving for each λ_t^*, the Chernoff information of all 512 pairs could be evaluated, leading to a ranking of the most different pairs (c_t^M, c_t^N), and the most discriminative k-vectors of **U**. The selected canonical vectors from **U** build new matrix $\widetilde{\mathbf{U}}$. In additional, the top k different pairs from (c_t^M, c_t^N) are selected.

Image Retrieval via Canonical Correlation Analysis

- Similarity measurement with hypothesis testing
  - Step 1 – Feature transformation:

    Given any two feature vectors (x_r , y_c), the transformed feature column vectors are computed as follows:

$$w = [w_1, w_2, \ldots, wk]^T = \widetilde{\mathbf{U}}^T \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}} x_r ,$$

$$v = [v_1, v_2, \ldots, vk]^T = \widetilde{\mathbf{U}}^T \boldsymbol{\Sigma}_{auto}^{-\frac{1}{2}} y_c .$$

Image Retrieval via Canonical Correlation Analysis

- Similarity measurement with hypothesis testing
  - Step 2 – Hypothesis testing:

We assume (x_r, y_c) comes from jointly gaussian distribution. Hence, (w_i, v_i) for $i \in$ {1,2,...,k} are jointly Gaussian, and independent from each other.

$$P_M(\mathrm{w_i}, \mathrm{v_i}) = \frac{e^{-\frac{1}{2}[\mathrm{w_i} \quad \mathrm{v_i}]\begin{bmatrix} 1 & \tilde{c}_i^M \\ \tilde{c}_i^M & 1 \end{bmatrix}\begin{bmatrix} \mathrm{w_i} \\ \mathrm{v_i} \end{bmatrix}}}{\sqrt{(2\pi)^2 \begin{vmatrix} 1 & \tilde{c}_i^M \\ \tilde{c}_i^M & 1 \end{vmatrix}}}, \quad P_N(\mathrm{w_i}, \mathrm{v_i}) = \frac{e^{-\frac{1}{2}[\mathrm{w_i} \quad \mathrm{v_i}]\begin{bmatrix} 1 & \tilde{c}_i^N \\ \tilde{c}_i^N & 1 \end{bmatrix}\begin{bmatrix} \mathrm{w_i} \\ \mathrm{v_i} \end{bmatrix}}}{\sqrt{(2\pi)^2 \begin{vmatrix} 1 & \tilde{c}_i^N \\ \tilde{c}_i^N & 1 \end{vmatrix}}}.$$

Image Retrieval via Canonical Correlation Analysis

- Similarity measurement with hypothesis testing
  - Step 2 – Hypothesis testing:

$$\text{Confidence score} = \sum_{i=1}^{k} \log \frac{P_M(\mathrm{w_i}, \mathrm{v_i})}{P_N(\mathrm{w_i}, \mathrm{v_i})}.$$

The higher the score is, the more likely the two images are a match. And we rank the confidence scores in descending order to obtain the image retrieval results.

Evaluation Results from 30k-SfM learning database on Oxford5k

Learning dataset: 30k-SfM													
	Dim	MAC				SPoC				SD			
		LDA	PCAw	S-CCA	G-CCA	LDA	PCAw	S-CCA	G-CCA	LDA	PCAw	S-CCA	G-CCA
Oxford5k	25	—	0.3504	0.2424	**0.3901**	—	0.4796	0.2511	**0.4879**	—	0.4993	0.3355	**0.5008**
	50	—	0.4264	0.3290	**0.4690**	—	0.5153	0.3149	**0.5437**	—	0.5129	0.4542	**0.5856**
	100	—	0.4980	0.4106	**0.5064**	—	0.5217	0.4549	**0.6219**	—	0.6038	0.5292	**0.6300**
	200	—	0.5547	0.4933	**0.5592**	—	0.6072	0.5123	**0.6658**	—	0.6580	0.6838	**0.6877**
	300	—	**0.5710**	0.5400	0.5406	—	0.6433	0.5274	**0.6723**	—	0.6728	0.6610	**0.6824**
	400	—	**0.5726**	0.5614	0.5463	—	0.6516	0.5373	**0.6713**	—	0.6825	0.6699	**0.6831**
	450	—	**0.5731**	0.5618	0.5424	—	0.6549	0.5333	**0.6696**	—	**0.6869**	0.6750	0.6812
	512	—	0.5620	**0.5621**	0.5418	—	0.6535	0.6535	**0.6704**	—	0.6805	0.6786	**0.6815**

Evaluation Results from 120k-SfM learning database on Oxford5k.

Learning dataset: 120k-SfM													
	Dim	MAC				SPoC				SD			
		LDA	PCAw	S-CCA	G-CCA	LDA	PCAw	S-CCA	G-CCA	LDA	PCAw	S-CCA	G-CCA
Oxford5k	25	0.3603	0.3830	0.2682	**0.4235**	0.4758	0.4472	0.3203	**0.4783**	0.4759	0.4779	0.3262	**0.5017**
	50	0.4760	0.4277	0.3720	**0.4780**	0.5612	0.4930	0.4085	**0.5627**	0.5375	0.5129	0.4521	**0.5688**
	100	0.5157	0.5185	0.4510	**0.5432**	0.6017	0.5675	0.5379	**0.6338**	0.6429	0.6038	0.5547	**0.6597**
	200	0.5887	0.5443	0.5516	**0.6182**	0.6571	0.6399	0.6440	**0.6947**	0.6861	0.6337	0.6485	**0.7176**
	300	0.6028	0.5619	0.5723	**0.6246**	0.6643	0.6575	0.6651	**0.7089**	0.7030	0.6638	0.6770	**0.7325**
	400	0.5974	0.5793	0.5680	**0.6251**	0.6688	0.6808	0.6699	**0.7116**	0.7020	0.6933	0.6824	**0.7381**
	450	0.5939	0.5819	0.5777	**0.6233**	0.6678	0.6862	0.6754	**0.7124**	0.6972	0.6952	0.6894	**0.7391**
	512	0.5868	0.5765	0.5765	**0.6229**	0.6613	0.6851	0.6851	**0.7131**	0.6958	0.6900	0.6900	**0.7393**

Image Retrieval via Canonical Correlation Analysis

Thank you ! Q&A…