P1-6

Patent Image Retrieval Using Cross-entropy-based Metric Learning

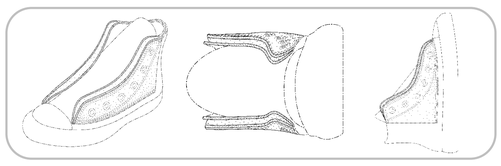
Kotaro Higuchi Yuma Honbu Keiji Yanai
The University of Electro-Communications, Tokyo, Japan

Introduction

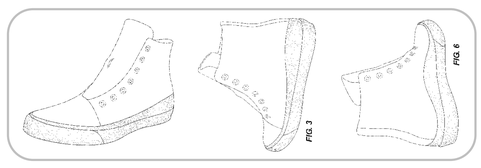
Backgrounds

1. Patents have a long challenge for image retrieval
   - Described as black-and-white drawings
2. De facto method or system has yet to emerge
3. Drawings are the focus of infringement judgments

➢ Yet to be a successful application in the past



USD811075S1



USD799800S1

Contributions

- Achieve SOTA score with metric learning
  - Introduce InfoNCE and ArcFace
- Implement a patent image retrieval application

Related work

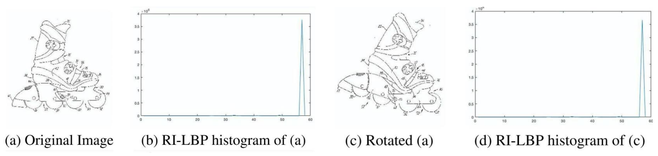
Conventional Patent Searches

➢ Search tools are keyword-based

Related work on patent image retrieval

- Feature descriptor[1]
  - Rotation invariant local binary patterns(RI-LBP)
- Dual VGG Net for patent image retrieval[2]
  - Trained as a classification task
- DeepPatent[3]
  - ResNet with training Triplet network

➢ We introduce more sophisticated metric learning methods.

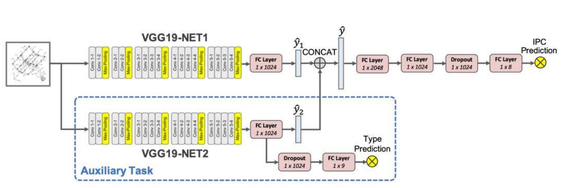


(a) Original Image (b) RI-LBP histogram of (a) (c) Rotated (a) (d) RI-LBP histogram of (c)

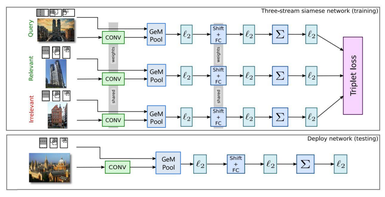
RI-LBP, feature descriptor[1]

$$\mathbf{f}^{(g)} = [\mathrm{f}_1^{(g)} \ldots \mathrm{f}_k^{(g)} \ldots \mathrm{f}_K^{(g)}]^\top$$
$$\mathrm{f}_k^{(g)} = \left(\frac{1}{|\chi_k|} \sum_{x \in \chi_k} x^{p_k} \right)^{\frac{1}{p_k}}$$

GeM Pooling[3]



Dual-VGG, classification task[2]



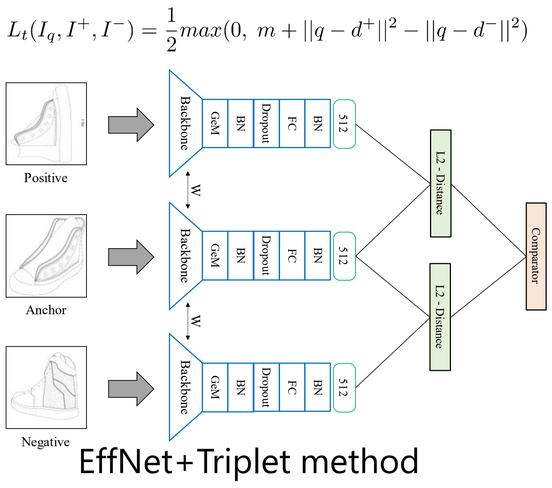
DeepPatent, triplet[3]

Method for patent image retrieval

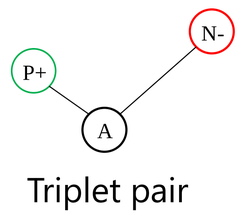
Baseline: Triplet[4]

- Triplet can handle one positive and one negative pair
  - Hard negative mining for Triplet training
  - Use EfficientNet as backbone

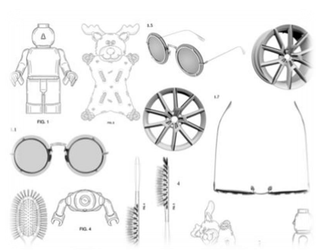
$$L_t(I_q, I^+, I^-) = \frac{1}{2} max(0,\ m + ||q - d^+||^2 - ||q - d^-||^2)$$



EffNet+Triplet method

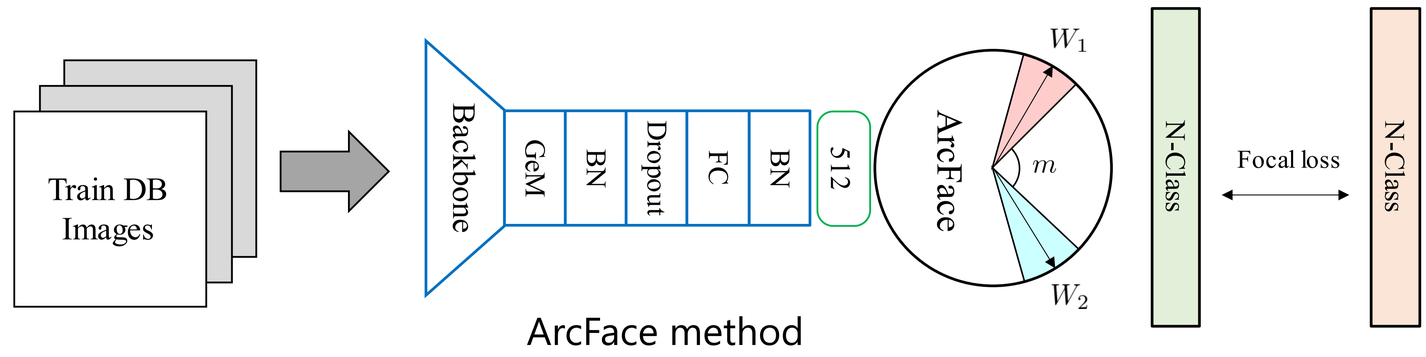


Triplet pair



DeepPatent samples[3]

Methods in our work



ArcFace method

InfoNCE[5]

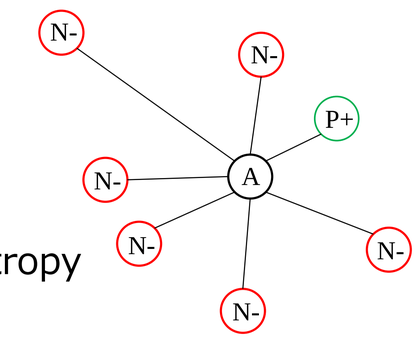
- InfoNCE uses many Anchor-Negative Pairs

$$L_i = -\log \frac{e^{q \cdot k_+ / \tau}}{e^{q \cdot k_+ / \tau} + \sum_{i=0}^{K} e^{q \cdot k_i / \tau}}$$

ArcFace[6]

- Add margin m for the correct class to the Cross-Entropy

$$L_a = -\log \frac{e^{s \cos(\theta_{y_i} + m)}}{e^{s \cos(\theta_{y_i} + m)} + \sum_{j=1, j \neq y_i}^{N} e^{s \cos \theta_j}}$$



infoNCE pair

Experimental results

mAP score comparison

Method	mAP
RI-LBP [3]	0.069
ResNet + Triplet [3]	0.379
EffNet + Triplet	0.384
EffNet + InfoNCE (ours)	0.447
EffNet + ArcFace (ours)	**0.662**

Dataset distribution

	Figures	Classes
Train	254,787	33,364
Test	38,834	6,927
Validation	44,815	5,888

Dataset: DeepPatent[3]

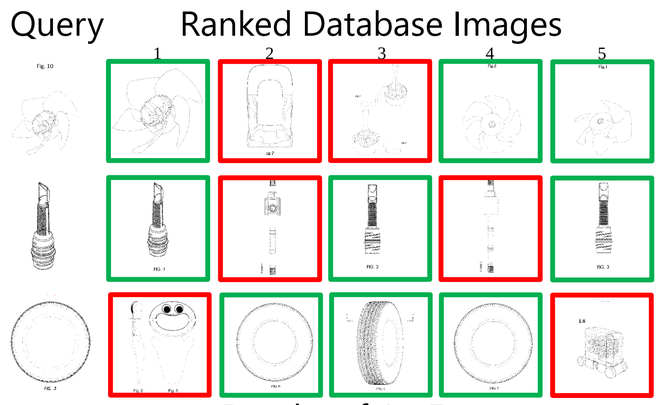
- Public dataset of the U.S. design patent
  The dataset size is 338k patent drawings

Evaluation Index

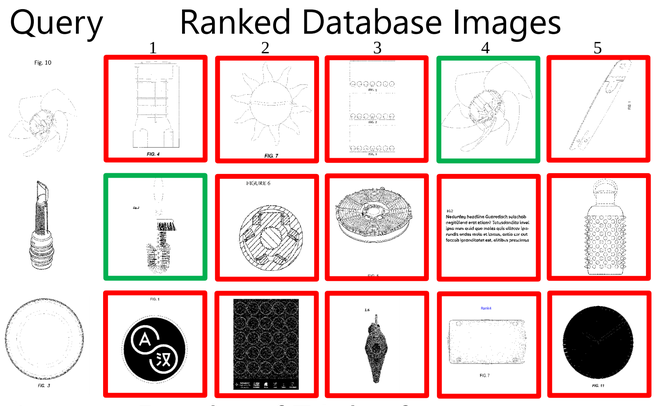
- mAP score: average precisions from each test query of the dataset
  Data split is the same as in the paper[3]

Comparison methods to baseline

➢ ArcFace achieved the best score

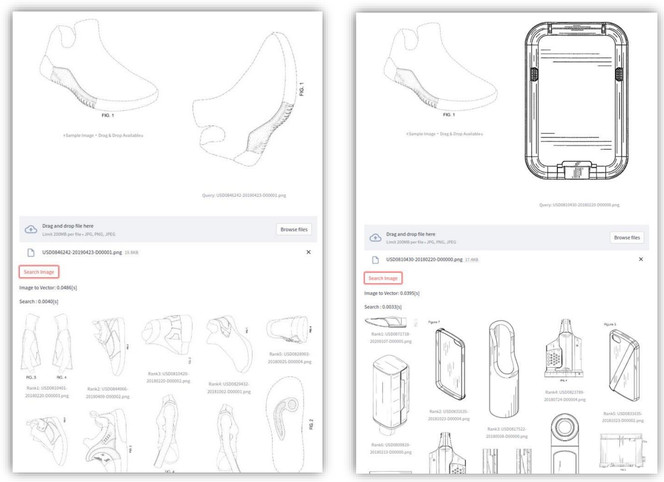


Results of ArcFace (Green: OK, Red: NG)



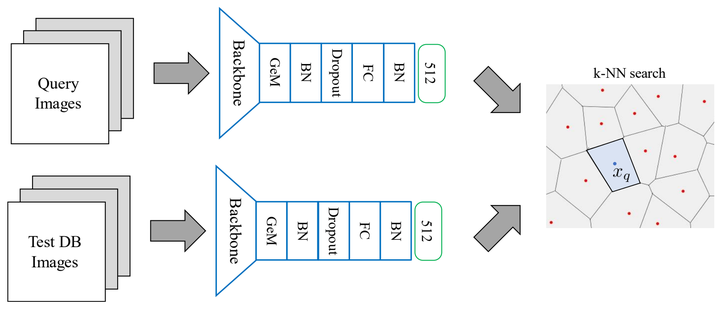
Results of Triplet for same query

Search application



(a) Shoes (b) Smartphone cases

- **Implementation using FAISS[7]**
  - Fast similarity search library, FAISS
  - ArcFace model on Ubuntu server



Retrieval/Inference architecture

[1] Yang, L., Gong, M., and Asari, V.: Diagram Image Retrieval and Analysis: Challenges and Opportunities. In: CVPR Workshops. pp.685-698 (2020)
[2] Jiang, S., Luo, J., Pava, G., Hu, J., and Magee, C.: A Convolutional Neural Network-based Patent Image Retrieval Method for Design Ideation. In: IDETC-CIE. ASME (2020)
[3] Kucer, M., Oyen, D., Castorena, J., and Wu, J.: DeepPatent: Large Scale Patent Drawing Recognition and Retrieval. In: WACV. pp. 2309-2318 (2022)
[4] Hoffer, E., and Ailon, N.: Deep Metric Learning Using Triplet Network. In: Similarity-Based Pattern Recognition. pp. 84-92. Springer International Publishing (2015)
[5] Oord, A., Li, Y., and Vinyals, O.: Representation Learning with Contrastive Predictive Coding. arXiv preprint arXiv:1807.03748 (2018)
[6] Deng, J., Guo, J., Xue, N., and Zafeiriou, S.: ArcFace: Additive Angular Margin Loss for Deep Face Recognition. In: CVPR. pp. 4685-4694 (2019)
[7] Facebook AI Similarity Search Library(FAISS): https://ai.facebook.com/tools/faiss/