Paper ID: #704

PS T5.5

AdaFilter: Adaptive Filter Design with Local Image Basis Decomposition for Optimizing Image Recognition Preprocessing

Aiga Suzuki[*†‡], Keiichi Ito[*], Takahide Ibe[*], Nobuyuki Otsu[‡]

[*]Adacotech Inc., [†]University of Tsukuba, [‡]National Institute of Advanced Science and Technology, Japan

Adacotech
Adaptive Cognition Technology

2020-01-14
@ICPR'20

Background

Importance of Image Preprocessing filter design in Image Recognition

Conventional: Finding the combination of typical image filtering

- Just a combination have lack of representability for complex features
- To obtain *"optimal filter"*, the model **must design** a filter **in pixel-level**

Our goal

To **design the optimal filter**, which maximize the expected generalization performance **for any task and any models,** as **not just a combination of typical filters** but ***in pixel-level***

The problem is the maximization of the generalization performance R:

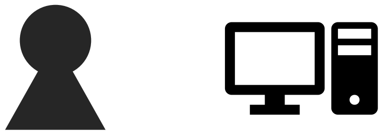
$$\underset{f}{\text{maximize}}\ R(f)$$

[Research problem]

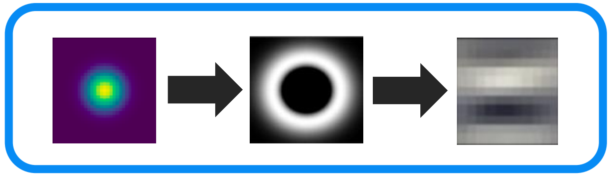
Designing f is in the $c \times k^2$-dimemsional space

Conventional manner

Target dataset



Engineer or Algorithm



Finding the combination

e.g.

(Blur=>Edge enhance=>Edge Extraction)

Adacotech
Adaptive Cognition Technology

Core Ideas

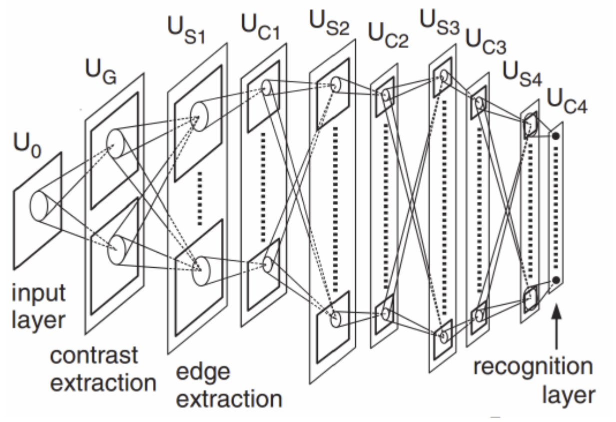
Represent an optimal filter as a linear combination of image basis

- The *linear filtering* is inherently **emphasizing/inhibiting the specific local patterns**

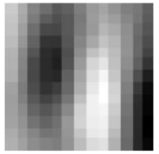
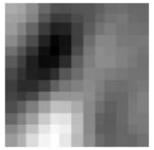
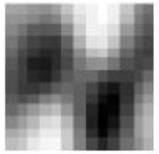
  - The core idea of *Neocognitron*, a.k.a. the origin of modern CNNs

- How can we represent **essential local patterns** *effectively*?

  - According to the theory of image reconstruction and sparse modeling, the **local image basis** could work as a *"good image atom"*



Neocognitron [Fukushima 1980]

0.4 ∗  + 0.2 ∗  - 0.7 ∗  + 0.6 ∗  = 

Linear combination of extracted local image bases

Complex Optimal Filter

Methodology: Generating Filter Bases

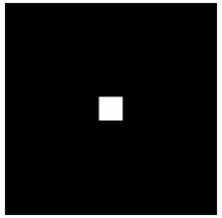
Extract a Filter basis from sampled local regions of training dataset

- Filter bases is **calculated as decomposed components** of sampled local region
  - Scan all $k \times k$ local regions from the training dataset to design $k \times k$ filter

Filter basis (atom) V

- Filter bases is **calculated as decomposed components** of sampled local region
  - The problem is formulated as a **Low-rank approximation** for sampled images
  - This study adopted **an Independent Component Analysis (ICA)** and **a sparse PCA (sPCA)** algorithms to calculate image bases

- To support a **representability of identity preprocessing** (*do nothing*), the **impulse filter is added** to the filter bases



impulse filter, which do nothing

Methodology: Maximizing Expected Performance

Maximize the Expected generalization performance as Black-Box Optimization

- By using filter bases V, the problem can be reformulated as **a coefficient optimization**

$$\underset{f}{\text{maximize}}\ R(f)$$

f: filter kernel

(in the $c \cdot k^2$-dim. space)

Relaxation $\Rightarrow$

$$\underset{\boldsymbol{a}}{\text{maximize}}\ R(V\boldsymbol{a})\ sbj.to\ \boldsymbol{a} \in [-1, 1]^M$$

V: filter bases, $\boldsymbol{a}$: coefficient vector

(in the M-dim. space)

M is the number of bases (hyper parameter)

- Expected generalization performance R **is approximated by validation data** *(hold-out, K-fold CV)*

- This is the Black-Box optimization, solved by search-based optimization algorithms
  - This study adopted **Bayesian Optimization (TPE)** and **Evolution Strategy (CMA-ES)** to solve this maximization problem

Adacotech
Adaptive Cognition Technology

Experiment

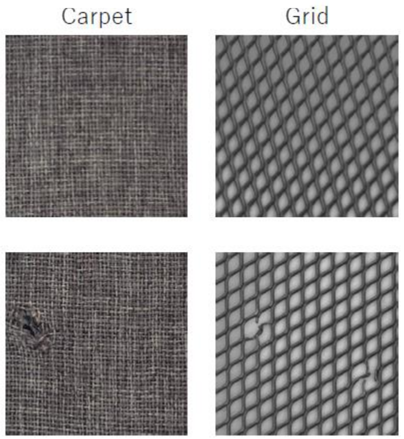
- In the experiment, we applied the proposed method to
  1. **Conventional ML algorithm-based anomaly detection**
  2. **CNN-based image classification task**
- We compared the generalization performance to **typical preprocessing filters**

1. ML-based Anomaly detection

Task: MVTec AD dataset

Method: HLAC image feature
+ Subspace Method

Metric: Validation ROC-AUC



** The case that preprocessing works dominantly*

2. DL-based Classification

Task: Caltech-101 dataset

Method: ResNet-50 CNN-feature
+ Linear-SVM

Metric: Validation F-measure

** CNNs are optimized for raw input*

Result

1. Anomaly detection for MVTec AD with HLAC+SM

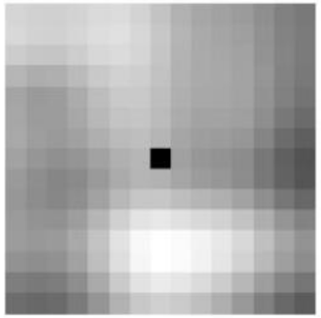
ROCAUC Task	Best score on typical	ICA+ES	sPCA+ES	ICA+TPE	sPCA+TPE
Carpet	*0.648*	**0.843**	0.717	0.816	0.722
Grid	*0.732*	0.908	**0.913**	0.891	0.901
Leather	*0.916*	0.965	0.954	**0.971**	0.969

2. Classification for Caltech-101 with CNN-feature + Linear-SVM

F1-score	Non-filtered	ICA+ES	sPCA+ES	ICA+TPE	sPCA+TPE
Carpet	0.865	**0.875**	*0.839*	**0.882**	*0.832*

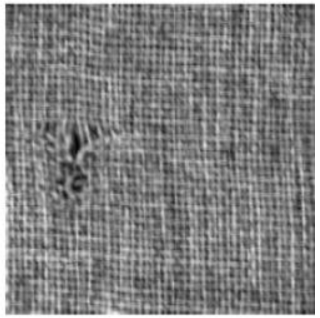
Result for MVTecAD (Carpet class)

Designed filter



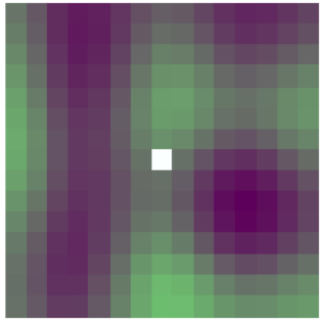
Original Image

Filtered Image



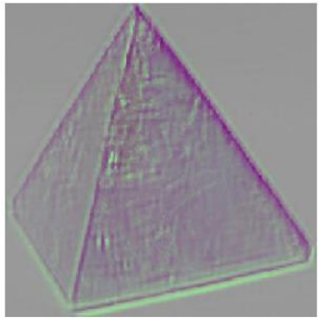
Result for Caltech-101 (whole dataset)

Designed filter



Original Image

Filtered Image



Adacotech
Adaptive Cognition Technology