# RSAC: Regularized Subspace Approximation Classifier for Lightweight Continuous Learning

Chih-Hsing Ho, Shang-Ho (Lawrence) Tsai
National Chiao Tung University

## Introduction

- Deep networks achieves a great success under the setting of supervised.
- Human possess the ability to continually grow the knowledge throughout the lifespan by solving different tasks.
- However, deep network fails to solve the old tasks when a new task is learned.
- This phenomenon is referred as catastrophic forgetting, where the classifier forgets the knowledge previously established after training on new data.
- The proposed approaches can be mainly categorized into weight consolidation, architecture expansion and memory rehearsal.
- Weight consolidation based methods often suffer from insufficient learning capacity as the flexibility is restricted by the regularization imposed for consolidating the old knowledge.
- Architecture expansion approaches are difficult to scale up in general when new coming tasks increase dramatically.
- Rehearsal based approaches store past examples which requires extra memory usage.
- Therefore, these prior works do not meet the requirement of lightweight continuous learning scenario (LCL).
- Inspired from SAAK [1], we proposed regularized subspace approximation classification (RSAC) for LCL. RSAC is a feedforward network as SAAK.

## Proposed method

- There are two modules in RSAC, feature reduction module and classifier module.
- $x$ is the raw image, $\mu_c$ is the mean of images and $\Sigma_c$ is the covariance, which belongs to class $c$.
- The equations is defined as,

$$\mu_c = \frac{1}{N_c}\sum_{j=1}^{N_c} x_j,\ \Sigma_c = \frac{1}{N_c}\sum_{j=1}^{N_c}(x_j-\mu_c)(x_j-\mu_c)^T = Q_c\Lambda_c Q_c^T$$

where $N_c$ is the number of data belongs to class $c$, and $\Lambda_c$ is a diagonal matrix with eigenvalues of class $c$ as its entry.

- Corresponding to the entry of $\Lambda_c$, $Q_c$ is a $d \times d$ orthonormal matrix composed of eigenvectors.
- We leverage KLT transform for feature reduction in lightweight continuous learning.

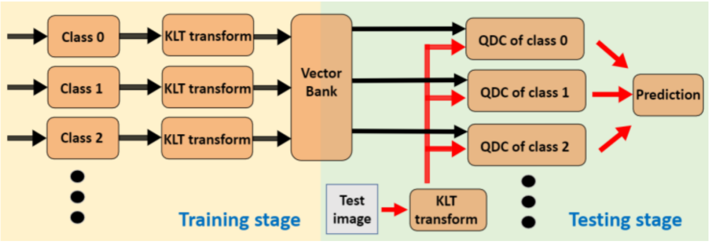


Figure 1. Regularized Subspace Approximation Classifier(RSAC) architecture

- The selecting of top k largest eigenvalues in $\Lambda_c$ is evaluated with the power threshold as

$$\arg\max_c \ln\left(\mathrm{P}_{Y|X}(c|f(x))\right) = \arg\max_c \ln\left(\mathrm{P}_{X|Y}(f(x)|c)\right) + \ln(\mathrm{P}_Y(c))$$

- Then, we project the input image $x$ on the top $k$ eigenvectors of $Q_c$, which is represented as $\hat{Q}_c$.

$$f(x) = \hat{Q}_c^T x$$

- In classifier module, it computes the latent representation f(x) through the maximum a posteriori in log scale to obtain final prediction.

$$\frac{\sum_{j=1}^{k}\sigma_c^j}{\sum_{j=1}^{d}\sigma_c^j} \geq t$$

- However, computing the inverse of covariance in latent representation is an ill-defined problem, we proposed to add a regularization on the covariance in latent representation.

$$\hat{\Sigma}_c = Cov(f(x)) = \hat{Q}_c^T Cov(x)\hat{Q}_c = \hat{Q}_c^T\Sigma_c\hat{Q}_c = \hat{Q}_c^T Q_c\Lambda_c Q_c^T\hat{Q}_c = \hat{\Lambda}_c$$

$$\hat{\Sigma}_c' = \hat{\Sigma}_c + \alpha * I = \hat{\Lambda}_c + \alpha * I$$

## Experiments

| Methods | Datasets (Accuracy) Mnist | KMnist | Fashion Mnist | Datasets (Training Time (sec)) Mnist | KMnist | Fashion Mnist |
|---|---|---|---|---|---|---|
| DGR [28] | 90.44±1.56 | 69.25±2.94 | 74.83±5.50 | 315.99±2.25 | 748.75±51.17 | 760.21±21.72 |
| DGR+distill [20], [28] | 92.31±0.74 | 64.42±1.12 | 76.03±4.12 | 314.12±12.79 | 819.52±14.52 | 800.81±3.69 |
| EWC [21] | 20.45±1.15 | 19.54±0.12 | 19.97±0.02 | 398.86±11.04 | 719.89±21.95 | 697.24±53.39 |
| Online EWC [61] | 20.69±1.53 | 19.54±0.12 | 19.97±0.03 | 371.87±12.35 | 665.04±3.40 | 692.49±29.20 |
| iCaRL [13] | 93.24±0.70 | 70.83±2.78 | 79.61±0.79 | 200.16±9.83 | 468.38±4.98 | 466.60±11.09 |
| LwF [20] | 20.98±0.85 | 20.16±0.24 | 19.42±2.54 | 198.40±9.09 | 495.62±31.48 | 499.49±8.77 |
| RtF [46] | 93.75±1.28 | 66.16±3.06 | 74.11±4.82 | 253.37±9.22 | 639.66±25.56 | 678.42±34.04 |
| SI [22] | 19.85±0.10 | 19.53±0.09 | 19.97±0.02 | 194.16±87.6 | 503.72±5.15 | 498.37±3.28 |
| CNDPM [62] | 93.54±0.13 | 74.35±1.4 | 44.62±2.1 | > 3600 | > 3600 | > 3600 |
| Saak [32] | 95.21 | 76.25 | 73.51 | > 3000 | > 3000 | > 3000 |
| Ours | **95.59** | **77.35** | **80.32** | **5.90** | **5.72** | **5.48** |

Figure 2. Class incremental scenario

| Power threshold $t$ | Mnist $k$ | Mnist acc | KMnist $k$ | KMnist acc | Fashion Mnist $k$ | Fashion Mnist acc |
|---|---|---|---|---|---|---|
| 0.8 | 31 | 67.75 | 64 | 61.30 | 26 | 64.90 |
| 0.9 | 68 | 93.22 | 126 | 76.16 | 77 | 73.98 |
| 0.95 | 121 | 95.41 | 211 | 77.13 | 156 | 79.74 |
| 0.96 | 141 | 95.43 | 243 | 76.87 | 185 | 80.25 |
| 0.97 | 168 | 95.43 | 285 | 74.84 | 224 | 73.56 |
| 0.98 | 206 | 91.66 | 346 | 75.09 | 278 | 73.95 |
| Best | 150 | 95.59 | 192 | 77.35 | 183 | 80.32 |

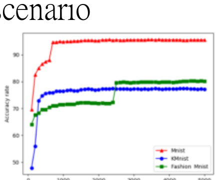


Figure 3. Ablation study (a) power threshold and (b) data incremental

## Reference

[1] C. Kuo and Yueru Chen. On data-driven saak transform.Journal ofVisual Communication and Image Representation, 50, 10 2017.