Very Low-Res Image Classification

Adam Vest[12], Muhammadabdullah Jamal[2], Boqing Gong[2]
[1]University of Louisville, [2]UCF Center for Research in Computer Vision (CRCV)

Abstract

The problem of classifying low-res (LR) images is both highly challenging as the information carried by each image is very limited and incredibly important for developing a robust, resolution-agnostic classification system. In the past researchers have attempted to solve the LR image classification problem directly, that is, by training classifiers on a dataset of low-res images. The difficulty of this problem can be seen in the relatively high classification error rates of state-of-the-art models on LR image datasets (e.g., the state-of-the-art CIFAR-100 error rate is nearly 5x that of the state-of-the-art ImageNet error rate [7]). In this work, however, we take a different approach and explore the use of a super-resolution network to transform input LR images to the corresponding high-res (HR) representations before classification. We show that using the above technique, we are able to decrease the error rate of LR image classification.

CONTACT

Adam Vest
UofL, CRCV
Email: apvest01@louisville.edu
Phone: (502) 594-0141

Problem Statement

- Low-res (LR) images contain limited information making recognition tasks such as classification more difficult [6]
- LR image classifiers trained from scratch on datasets with a non-trivial number of classes tend to have relatively high error rates [7]
- Real-world classification systems need to be flexible and handle any LR images they may encounter without suffering serious performance degradation
- Deep ConvNets and GANs [4] have been shown to outperform traditional SR algorithms such as Bicubic Interpolation (BC)



Fig. 1. Example results of several SR algorithms on a patch of an input image [2]

Hypotheses and Motivation

- Through the use of super-resolution, we may improve LR image classification performance
- The performance of SR-based classification systems should improve as the quality of SR images improves, so deep ConvNet and GAN approaches should outperform BC
- **Goal:** to develop a resolution-agnostic image classification system that utilizes super-resolution to improve LR image classification performance

Experiment Setup

- **Training details:**
  - Initial learning rate of 10^{-3}, decayed by a factor of 10 if validation accuracy fails to increase by .5% for 3 epochs
  - Training ended if during a period of 10 epochs the validation accuracy had not improved by .001%
  - All AlexNet [3] based models were initialized with ImageNet [1] pretrained weights
  - A reduced learning rate was used for fine-tuned layers
  - All models were implemented using Pytorch [8]
- **Dataset Description:**
  - Used Caltech101 dataset (50%, 25%, 25% train, validation, testing split)
  - Preprocessed images by scaling to 256, center-cropping to 224x224, then downscaling to 56x56
- **Performed Experiments:**
  - Simple AlexNet-based LR image classifier
  - HR image classification using (1) AlexNet as a feature extractor and (2) fine-tuning AlexNet
  - Classified LR images super-resolved by Bicubic Interpolation using (1) pretrained HR feature extractor AlexNet and (2) pretrained HR fine-tuned AlexNet
  - Classified LR images super-resolved by SRResNet [4] using (1) pretrained HR feature extractor AlexNet and (2) pretrained HR fine-tuned AlexNet

Results

Model	Top-1 Accuracy	Top-5 Accuracy
Simple LR Classifier	50.07%	70.80%
HR Images w/ Feature Extractor AlexNet	79.91%	93.55%
HR Images w/ Fine-tuned AlexNet	85.68%	96.57%
BC SR Images w/ Pretrained HR Feature Extractor Model	54.14%	80.87%
BC SR Images w/ Pretrained Fine-tuned HR Model	67.60%	88.24%
SRResNet SR Images w/ Pretrained HR Feature Extractor Model	70.62%	88.38%
SRResNet SR Images w/ Pretrained Fine-tuned HR Model	78.63%	92.68%

Table 1. Top-1 and Top-5 Classification Accuracies of HR Baseline Models and LR Classification Models

Model Diagrams

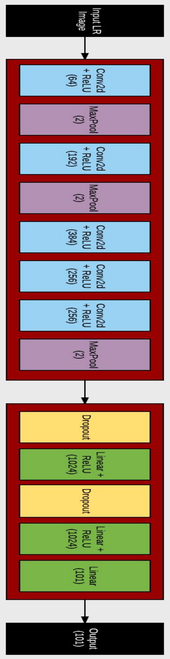


Fig. 2. Simple Classifier Architecture

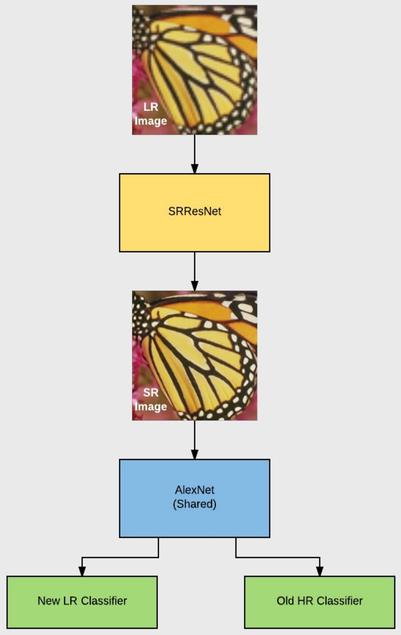


Fig. 3. Proposed LwF-based model for a robust resolution-agnostic classification system using Super-Resolution

Discussion

- Initial results support Hypotheses 1 and 2
- The models may be grouped based on classification accuracy (in descending order) as follows:
  - **HR Classifiers:** The combined benefits of large, readily available HR image datasets such as ImageNet and high quality train/test images likely account for these models' success
  - **SR Classifiers:** It is likely that these models outperform the LR classifiers because they can leverage large HR image datasets and are given train/test images with more discriminative details
  - **LR Classifiers:** These models likely perform poorly due to the limited information carried by LR images and the lack of a base model pretrained on a large dataset
- Models utilizing SRResNet to super-resolve LR images significantly outperformed models using Bicubic Interpolation, suggesting that SR image quality plays an important role in enabling pretrained HR classifiers to effectively classify LR images
- Significant performance boosts were observed when the base classifier (AlexNet) was fine-tuned rather than being used as a feature extractor

Future Work

- Work on Super-Resolution for Classification, i.e., a method of fine-tuning the SRResNet to produce better classification results
- Use the Learing without Forgetting [5] framework to create a multiple resolution image classification system
- Run models multiple times and determine the statistical significance of the observed results
- Explore the utility of super-resolution networks to classify very low-res images (i.e., even smaller images requiring higher upscaling factors to classify)

References

1. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In CVPR, 2009.
2. D. Glasner, S. Bagon, M. Irani. Super-resolution from a single image. In ICCV, 2009.
3. A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks. In NIPS, 2012.
4. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. ArXiv e-prints, September 2016.
5. Z. Li and D. Hoiem. Learning without forgetting. In ECCV, 2016.
6. Y. M. Lui, D. Bolme, B. A. Draper, J. R. Beveridge, G. Givens, and P. J. Phillips. A meta-analysis of face recognition covariates. In Proceedings of International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2009.
7. S. Xie, R. Girshick, P. Dollar, Z. Tu, and K. He. Aggregated residual transformations for deep neural networks. In CVPR, 2017.
8. Pytorch. http://pytorch.org/.

This work was support by NSF grant No. 100461121