10/26/2022

Literature Review and architecture

Final project deliverable 2

Group 15
Nourhan Abdelrhim
Alaa Khashap
Eman Aboud
Sondos Mohamed

Introduction and problem

Recently, most classification problems assumed that images have high resolution (HR) to extract features. In addition, most algorithms require a large amount of data to be able to learn. That's when transfer learning was introduced to allow low resolution (LR) and less amount of dataset to be learnt based on other learnt models. It is a model trained on a huge high-resolution dataset having the needed model parameters in the last 2 or 3 layers to learn them in addition. Since we try to solve a makeup classification problem with a low number of instances (500) and 10 classes with low resolution (64*64), we aim to apply multiple methods including transfer learning to classify our objects.

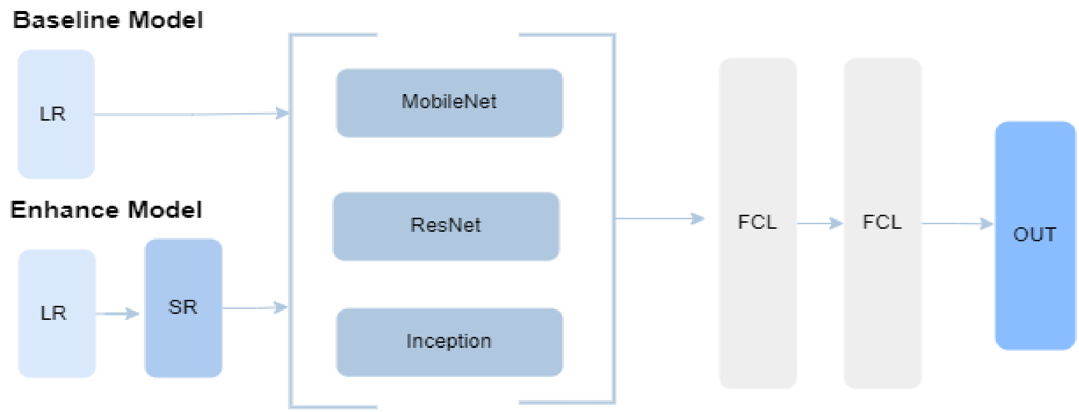
Literature review

Multiple research papers introduced solutions for low-resolution data classification. In this paper (Wu, Y., Zhang, Z. and Wang, G. (2019)), the authors have applied transfer learning to solve the LR problem without finetuning the convent filters, just extracting features from high-resolution and low-resolution images, then feeding them to a feature transfer network with 2 layers (shallow network), then using SVM for classification. They claimed that this network can be embedded into any other DNN for enhancement. This method starts with feature extraction of both LR and HR images using pool5 layer from ResNet-101 that was used as the backbone convnet, then unsupervised deep feature transfer that clusters the HR features and takes the LR features as testing to assign labels, then feed LR features to SVM classifier. The dataset is the same for LR and HR, just resizing was introduced. Grid search was used to get the optimal parameters. Baseline-HR and Baseline-LR were comparable baselines, and the proposed methodology was 2% higher in performance.

Another solution was presented by (Koziarski, M. and Cyganek, B. (2018)), they applied pre-trained models such as AlexNet, VGGNet, and ResNet to measure the classification accuracy of different neural architectures on images with varing degrees of low resolution. As a pre-processing method for low-resolution images, they used the image super-resolution (SR) using the VDSR network. Low image resolution's effects on different convolutional neural networks' classification accuracy. And they realize that in these situations, applying super-resolution before classification may be the most effective approach. They assessed VDSR, a trimming super-resolution technique, to determine its efficacy as a pre-processing technique for the image identification challenge. In addition, they held experiments to check whether super-resolution might be applied before classification to increase classification accuracy for low-resolution images. Finally, they introduced super-resolution into the image recognition pipeline to improve classification accuracy.

In addition, in trying to solve the low-resolution problem, The authors of this paper (Cai, D. et al. (2017)) suggested a unique resolution-aware deep model that integrates end-to-end convolutional fine-grained classification and convolutional image super-resolution into a single model. As convolutional super-resolution (SR) layers in resolution-aware convolutional neural network (RACNN), the authors use three conventionally stacked convolutional-ReLU layers with zero-padding filters, and they conducted experiments by using three commonly datasets to demonstrate the proposed RACNN's improved performance on low-resolution images when the performance of conventional CNN collapses. The results indicated that adding convolutional super-resolution layers to traditional CNNs can effectively improve performance in low-resolution fine-grained classification and recover fine features for low-resolution images. Kabir, M. M. et al. (2020) examined current CNN designs that perform better in terms of accuracy on small, poor-quality datasets. 500 photos from 10 classes, including mobile, pen, mouse, keyboard, speaker, paper, spray, book, pencil-box, and person, were utilized in the dataset. All 5000 photos were captured by web cameras and fed into the architecture. Six well-known pre-trained models were used including Densenet, Inception, Mobilenet, ResNet, VGG, and Xception. It might be considered that almost all models provide results with higher accuracy. Due to the training dataset's overfitting, the accuracy of ResNet and VGG architecture is about 0.9. MobileNet achieves the highest accuracy of 0.97 out of all the models. They note that well-known baseline architectures like Xception, DenseNet, ResNet, etc. perform admirably on high-quality ImageNet datasets but fall short on low-quality image datasets, failing to deliver the desired outcome. However, MobileNet operates best with low-resolution photos. Our

Architecture



Based on the literature review, we aim to apply this approach for classifying our objects. Mobile-Net, ResNet, and Inception pre-trained models will be used 2 times. The baseline model takes the Low-resolution images directly to the pre-trained models, then use the last 2 fully connected layers to save our parameters' values, and then the softmax dense output layer. The other enhanced solution includes the same steps but adds a super-resolution layer to the input. The number of layers will be tuned using grid search.

For more enhancement, we aim to try the high-resolution dataset that was collected firstly as the input layer in the same architecture, and after training, the model will be saved and used again as a pre-trained model (contains HR data trained on the main pre-trained models) and replace the input layer with the LR data

References

Wu, Y., Zhang, Z. and Wang, G. (2019) “Unsupervised deep feature transfer for low resolution image classification,” *arXiv [cs.CV]*. Available at: https://openaccess.thecvf.com/content_ICCVW_2019/papers/RLQ/Wu_Unsupervised_Deep_Feature_Transfer_for_Low_Resolution_Image_Classification_ICCVW_2019_paper.pdf (Accessed: October 25, 2022).

Koziarski, M. and Cyganek, B. (2018) “Impact of low resolution on image recognition with deep neural networks: An experimental study,” *International Journal of Applied Mathematics and Computer Science*, 28(4), pp. 735–744. doi: 10.2478/amcs-2018-0056.

Cai, D. *et al.* (2017) “Convolutional low-resolution fine-grained classification,” *arXiv [cs.CV]*. Available at: http://arxiv.org/abs/1703.05393 (Accessed: October 25, 2022).

Kabir, M. M. *et al.* (2020) “An evolution of CNN object classifiers on low-resolution images,” in *2020 IEEE 17th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET)*. IEEE, pp. 209–213.