Mar 19th, 9:00 AM - 11:00 AM

Two-Stage Transfer Learning for Facial Expression Classification in Children

Gregory Hubbard
Old Dominion University

Megan Witherow
Old Dominion University

Khan Iftekharuddin
Old Dominion University

OLD DOMINION UNIVERSITY

Two-Stage Transfer Learning for Facial Expression Classification in Children

Gregory Hubbard, Computer Science, May 2022

Co-authors: Megan A. Witherow, Khan M. Iftekharuddin (Faculty Mentor)

Abstract

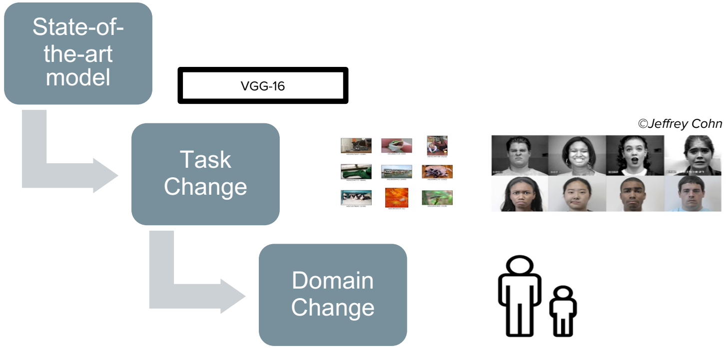
This work uses two transfer learning steps, a task change from general object classification to facial expression classification and a domain change from adult to child data, to train a neural network model for child facial expression classification.

Introduction & Literature

Transfer learning allows for knowledge transfer from a learned task to a new, related task. In this project, transfer learning is implemented to leverage a state-of-the-art pretrained model (general object recognition) on facial expression classification for children.

Methods

Transfer learning is implemented by freezing layers in the VGG-16 model. Two stages of transfer learning are used: general object recognition to adult facial expressions then adult expressions to child expressions.



Set	Loss	Accuracy
Training	0.0602	99.65%
Validation	0.7376	85.94%
Test	0.3981	88.73%

Email: ghubb002@odu.edu

Results

A test accuracy of 88.73% is achieved on an independent set of facial expression data from 10% of the child subjects. The most confusing classes for the model are anger and disgust.

Discussion

This work demonstrates the use of existing state of the art models (ex: VGG16) in two-stage transfer learning for child facial expression classification. To better understand the potential generalizability of the model, future work should include cross validation.

Works Cited

[1] Megan A. Witherow, Manar D. Samad, Khan M. Iftekharuddin, "Transfer learning approach to multiclass classification of child facial expressions," Proc. SPIE 11139, Applications of Machine Learning, 1113911 (6 September 2019); https://doi.org/10.1117/12.2530397

[2] Chollet, F., & others. (2015). Keras. GitHub. Retrieved from https://github.com/fchollet/keras

[3] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," CoRR, vol. abs/1409.1556, 2014.

[4] Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00), Grenoble, France, 46-53.

[5] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, 94-101.

[6] LoBue, V., & Thrasher, C. (2015). The Child Affective Facial Expression (CAFE) set: validity and reliability from untrained adults. Emotion Science, 5, 1532. http://doi.org/10.3389/fpsyg.2014.01532

[7] LoBue, V. (2014). The Child Affective Facial Expression (CAFE) set. Databrary. http://doi.org/10.17910/B7301K

Two-Stage Transfer Learning for Facial Expression Classification in Children: Abstract

Studying facial expressions can provide insight into the development of social skills in children and provide support to individuals with developmental disorders. In afflicted individuals, such as children with Autism Spectrum Disorder (ASD), atypical interpretations of facial expressions are well-documented. In computer vision, many popular and state-of-the-art deep learning architectures (VGG16, EfficientNet, ResNet, etc.) are readily available with pre-trained weights for general object recognition. Transfer learning utilizes these pre-trained models to improve generalization on a new task. In this project, transfer learning is implemented to leverage the pretrained model (general object recognition) on facial expression classification. Though this method, the base and middle layers are preserved to exploit the existing neural architecture. The investigated method begins with a base-packaged architecture trained on ImageNet. This foundation is then task changed from general object classification to facial expression classification in the first transfer learning step. The second transfer learning step performs a domain change from adult to child data. Finally, the trained network is evaluated on the child facial expression classification task.

Email: ghubb002@odu.edu

Two-Stage Transfer Learning for Facial Expression Classification in Children: Introduction

- Training a neural network is computationally expensive and takes a long amount of time.
- Transfer learning allows for knowledge transfer from a learned task to a new, related task through the importation of a pre-trained set of weights.
- To date, there are dozens of top-performing state-of-the-art neural network models for image recognition.
- Many of these models are available pre-trained on a general object recognition task and may be downloaded for use in transfer learning.
- In this project, transfer learning is implemented to leverage a state-of-the-art pretrained model (general object recognition) on facial expression classification for children.
- Since the child facial expression database is small, two stages of transfer learning are used: general object recognition to adult facial expressions and adult facial expressions to child facial expressions.

Email: ghubb002@odu.edu

Two-Stage Transfer Learning for Facial Expression Classification in Children: Literature

- Past transfer learning work for classification of facial expressions in children trains on adult expression data and then fine-tunes the network on child expression data [1].
- It is expected that the benefits of transfer learning will enhance and optimize the ability to characterize child facial affects.
- Convolutional neural networks (CNN) may take weeks or months to train. By implementing a pretrained network, the modified model takes advantage of already existing feature detections. Additionally, due to generalized training, a transfer learning model can function to better isolate specific features in training data from a new task.
- The Keras API provides immediate access to popular state-of-the-art models such as VGG (e.g. VGG16 or VGG19), GoogLeNet (e.g. InceptionV3), and Residual Network (e.g. ResNet50) [2].
- In the past, the VGG-16 model won the ImageNet challenge and achieved near state-of-the-art results in terms of prediction accuracy [3].

Email: ghubb002@odu.edu

Two-Stage Transfer Learning for Facial Expression Classification in Children: Data

Extended Cohn-Kanade (CK+) Database [4, 5]

- widely used for facial expression classification methods
- 593 video sequences
- 123 subjects, 18 to 50 years in age and representing a variety of genders and backgrounds
- Each video shows a facial shift from the neutral expression to a targeted peak expression, recorded at 30 frames per second
- Seven expressions : anger, contempt, disgust, fear, happiness, sadness, and surprise

Child Affective Facial Expression (CAFE) Set [6, 7]

- first and largest representation of juvenile facial expressions for scientific research
- 1192 images of juvenile facial expressions
- Subjects are between 2 and 8 years of age, and represent a diverse and wide range of ethnic / gender identities.
- Seven expressions: happy, angry, sad, fearful, surprise, neutral, and disgust

Fig. 1. Examples from the CK+ Database [4,5]

Email: ghubb002@odu.edu

Two-Stage Transfer Learning for Facial Expression Classification in Children: Method

- The VGG-16 model [3] is used as the base architecture.
- VGG-16 has three main parts: convolution, pooling, and fully connected layers.
- Following the base model, a flatten and dense layer are added to classify the expressions from the convolutional layers.
- The data are divided in a 90/10 train-test split. This means that 90% of subjects were used for training and 10% of the subjects were used for testing. 10% of the training data is used as a validation set for tuning the number of training epochs.
- Transfer learning is implemented by freezing model layers, i.e., not updating the weights in the layer during training. A reasonable number of frozen layers is selected via trial and error to maximize performance on the validation set.
- Transfer learning is done in two stages: general object recognition to adult facial expressions then adult expressions to child expressions.

```
flag = False

for layer in model1.layers[0].layers:
    layer.trainable = flag
    if layer.name == "block4_conv1":
        flag = True
    print(layer.name, layer.trainable)
```

Fig. 2. Code snippet for freezing model layers

```
input_1 False
block1_conv1 False
block1_conv2 False
block1_pool False
block2_conv1 False
block2_conv2 False
block2_pool False
block3_conv1 False
block3_conv2 False
block3_conv3 False
block3_pool False
block4_conv1 False
block4_conv2 True
block4_conv3 True
block4_pool True
block5_conv1 True
block5_conv2 True
block5_conv3 True
block5_pool True
Model: "sequential"
```

Layer (type)	Output Shape	Param #
vgg16 (Model)	(None, 8, 8, 512)	14714688
flatten (Flatten)	(None, 32768)	0
dense (Dense)	(None, 7)	229383

```
Total params: 14,944,071
Trainable params: 229,383
Non-trainable params: 14,714,688
```

Fig. 3. (top) VGG-16 model architecture showing trainable (True) and frozen (False) layers; (bottom) overall architecture

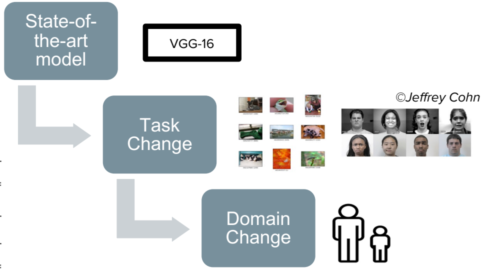


Fig. 4. Two-stage transfer learning pipeline

Email: ghubb002@odu.edu

Two-Stage Transfer Learning for Facial Expression Classification in Children: Results

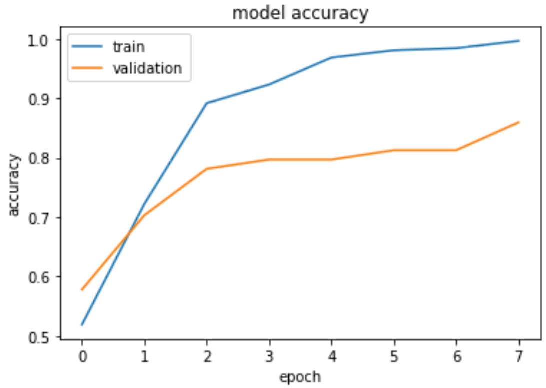


Fig. 5. Training and validation accuracy curves versus training epoch number

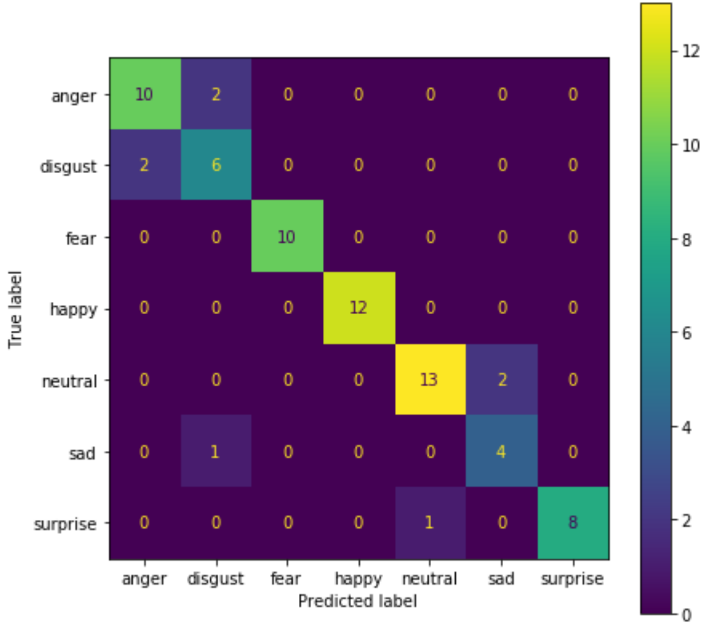


Fig. 6. Confusion matrix for CAFE [6, 7] test data

Table I. Training, Validation, and Test Loss and Accuracy Metrics

Set	Loss	Accuracy
Training	0.0602	99.65%
Validation	0.7376	85.94%
Test	0.3981	88.73%

Email: ghubb002@odu.edu

Two-Stage Transfer Learning for Facial Expression Classification in Children: Discussion

- This work demonstrates the use of existing state of the art models (ex: VGG16) in two-stage transfer learning for child facial expression classification.
- The final implementation considers the VGG16 layers frozen until the block4_conv1 layer. The layers beginning with block4_conv2 are then trained.
- The validation set is used to tune the number of epochs, resulting in training for 7 epochs.
- The gap of 13.71% between training and validation accuracies indicates overfitting.
- The most confusing classes for the model are anger and disgust, likely due to visual similarity.
- Some neutral expressions are confused as sad.
- A test accuracy of 88.73% is achieved on this train/test split.
- To better understand the potential generalizability of the model, future work should include cross validation.

Email: ghubb002@odu.edu

Two-Stage Transfer Learning for Facial Expression Classification in Children: Works Cited

[1] Megan A. Witherow, Manar D. Samad, Khan M. Iftekharuddin, "Transfer learning approach to multiclass classification of child facial expressions," Proc. SPIE 11139, Applications of Machine Learning, 1113911 (6 September 2019); https://doi.org/10.1117/12.2530397

[2] Chollet, F., & others. (2015). Keras. GitHub. Retrieved from https://github.com/fchollet/keras

[3] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” CoRR, vol. abs/1409.1556, 2014.

[4] Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00), Grenoble, France, 46-53.

[5] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, 94-101.

[6] LoBue, V., & Thrasher, C. (2015). The Child Affective Facial Expression (CAFE) set: validity and reliability from untrained adults. Emotion Science, 5, 1532. http://doi.org/10.3389/fpsyg.2014.01532

[7] LoBue, V. (2014). The Child Affective Facial Expression (CAFE) set. Databrary. http://doi.org/10.17910/B7301K

Email: ghubb002@odu.edu

Two-Stage Transfer Learning for Facial Expression Classification in Children: Acknowledgements

This work is partially supported by the National Science Foundation under Grant No. 1753793 and Grant No. 1828593.

VISION LAB
@ OLD DOMINION UNIVERSITY

Computer Vision • Machine Learning • Medical Image Processing

Email: ghubb002@odu.edu