CNN-Based Emotion Classification in Visual Art for Therapeutic and Creative Applications

Valentina Franzoni, Zofia Ruta

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF PERUGIA, ITALY, VALENTINA.FRANZONI@UNIPG.IT

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

ADAM MICKIEWICZ UNIVERSITY, POZNAM, POLAND

Emotion Recognition From artworks

- Enhance experience in art shows
- Improve UX with synchronised lights, music, multimedia
- Art therapy

Convolutional Neural Networks (CNN)

- CNNs + Transfer Learning

- VGG16
- InceptionV3
- MobileNetV2

- Prototype application

The relevance of data

WikiArt Emotion

+

ArtEmis

=new dataset

DISTRIBUTION OF EMOTION CLASSES AFTER MERGING WIKIART EMOTION AND ARTEMIS DATASETS

Emotion	Number of Samples	Percentage (%)
Anger	438	10.63%
Disgust	700	16.99%
Fear	567	13.76%
Happiness	1044	25.34%
Sadness	637	15.46%
Surprise	734	17.82%
Total	**4120**	**100%**

Examples of original data

Contentment

"the women are enjoying their time together"

Sadness

"the woman looks like she is being forced to be in a hurry"

ArtEmis

Annotations:

Avg. art rating: 2.25

Emotions in image: happiness, trust

WikiArt Emotion

Training and optimization

- Adam optimizer
  - Learning rate: 0.0001
  - Batch size: 32
- Training for 50 epochs
- Cross-entropy loss
- 80/20% data split for training/test
- Accuracy, precision, recall, and F1

Results (multiclass)

InceptionV3 outperformed other classifiers.

COMPARISON OF DEEP LEARNING MODEL PERFORMANCE (VGG16, MOBILENETV2, AND INCEPTIONV3) ON EMOTION CLASSIFICATION TASKS USING TRANSFER LEARNING. RESULTS SHOW INCEPTIONV3 ACHIEVING THE HIGHEST ACCURACY. IN BOLD, THE BEST RESULT FOR EACH CLASSIFIER.

Classifier	Optimizer	0.001	0.01
3*InceptionV3	adam	41.26%	41.38%
	rmsprop	39.56%	36.29%
	sgd	40.05%	**44.54**%
3*MobileNetV2	adam	21.60%	16.75%
	rmsprop	28.76%	13.96%
	sgd	37.99%	**40.78**%
3*VGG16	adam	41.88%	41.38%
	rmsprop	**42.11**%	40.53%
	sgd	32.77%	33.86%

Results (binary)

InceptionV3 outperformed other classifiers.

CONFUSION MATRIX FOR BINARY CLASSIFICATION USING THE INCEPTIONV3 MODEL. IT ILLUSTRATES THE CLASSIFICATION ACCURACY OF PLEASANT AND UNPLEASANT EMOTIONS, HIGHLIGHTING COMMON MISCLASSIFICATIONS.

	Predicted Pleasant	Predicted Unpleasant
Actual Pleasant	78.98%	21.02%
Actual Unpleasant	27.11%	72.89%

Unpleasant=*anger, disgust, sadness, fear (2342 samples);*

Pleasant=*happiness, surprise (1778 samples)*

Conclusions

- CNNs can be effective for ER in visual art
- Combining multiple datasets is beneficial
- Binary classification has better performance
- InceptionV3 performs reasonably well

Future improvements

- Diverse data for complex emotions
- Data augmentation to improve generalization
- Multimodal (visual + textual) models

Contacts

For informations or Q&A, write to valentina.franzoni@unipg.it

Find us on Facebook @affectivecomputing

www.facebook.com/affectivecomputing