Exploring Automatic Translation Between Affect Representation Schemes: Image Content Analysis

Author: Shuang Liu (s.liu-22@student.tudelft.nl)
Supervisor: Bernd Dudzik
Responsible Professor: Chirag Raman

1 Background

- **Images** as a medium is very powerful and can convey rich affect.
- **Image analysis** is the extraction of meaningful information from images[1]. If the information is about affect, then image content analysis can be seen as an affect prediction problem.
- The **representation** of affective states is a crucial component of affect prediction systems. It determines how the system understands and responds to affect.
  - **Categorical emotion states (CES)**
  - **Dimensional emotion space (DES)**
  - **CES: Ekman** (anger, disgust, fear, joy, sadness, surprise)
  - **DES: VA(D)** (valence, arousal, dominance)

2 Research Questions

- Which **Machine Learning model** performs best in the translation from DES to CES?
- What **factors** could influence the capacity of models to generalize to unseen datasets?

3 Method

Database Collection

- **Emotion6**[2] : Ekman+neutral, VA (SAM)

Images from Flickr, a popular online image-sharing platform. All of the images are put on Amazon Mechanical Turk (AMT) to be labeled with emotional keywords and annotated with VA scores

- **IAPS (Libkuman)**[3] : Ekman , VA (SAM)

The participants involved in the process of annotating images of the other dataset are 1,302 Midwestern university students who were 18 years old or older and included both males and females

Database Combination

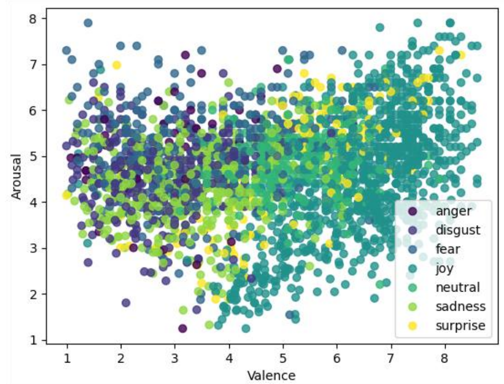

Figure1: The distribution of VA scores of Emotion6 and IAPS

Table1: Number of each emotion category of Emotion6 and IAPS

Emotions	Emotion6	IAPS	Total
anger	31	22	53
disgust	245	68	313
fear	329	68	397
joy	638	373	1,011
neutral	325	0	325
sadness	308	107	415
surprise	104	64	168

- An **imbalance** in the sample distribution across different emotion categories
- **Synthetic Minority Oversampling Technique (SMOTE)**: the number of each emotional state became 1,011

- **Overlap** among the emotion categories exists
- Considering both VA dimensions and categorical emotion labels when studying emotions is important

Model Implementation

- **Models**: KNN, DT, NB
- **Baseline**: Majority Classifier (MC)

Hyperparameters Identification

- **GridSearchCV**

DT
- max_depth: 2,4,8,16,32
- min_samples_leaf: 2,4,8,16

KNN
- n_neighbors: 2,5,10,15
- algorithm: ball_tree, kd_tree, auto
- metric: minkowski, euclidean, manhattan, chebyshev

NB
- var_smoothing: np.logspace(0,-9,num=100)

Figure2: Types and ranges of hyperparameters of each model

Evaluation Metric

- **Accuracy** (Acc)

Model Selection

- **30 trials** in total
- **Nested cross-validation** to select a model
  Inner & outer loop: 5-fold cross-validation
- **Inner loop**: Hyperparameters are tuned
- **Outer loop**: The model's performance on independent test sets are assessed

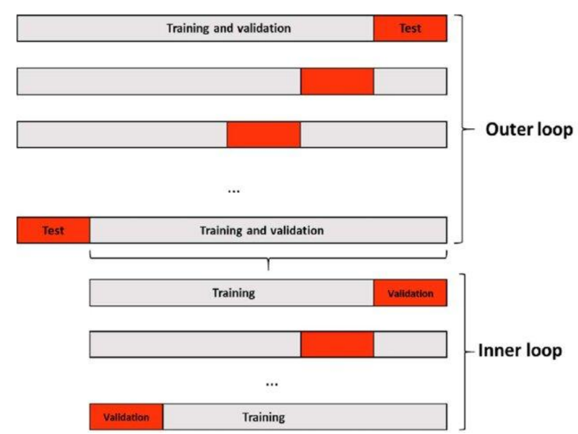

Figure3: Overview of nested cross-validation[4]

4 Results

Table2: Summary performance of each ML model, compared with the majority classifier

Models	Acc M	Acc SD	vs.Majority Δ M(Acc)	vs.Majority t	vs.Majority p
DT	0.586	0.005	+0.192	431	<.001***
KNN	0.610	0.004	+0.191	617	<.001***
NB	0.386	0.001	+0.201	581	<.001***
MC	0.131	0.002	0	0	1

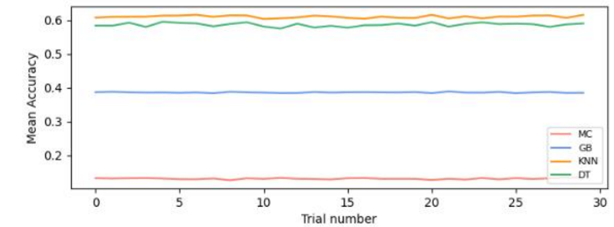

Figure4: A graphical overview of the performance of each model

Table3: T-test scores and p-values among classifiers

Models	t	p
KNN-DT	20	<.001***
DT-NB	196	<.001***
KNN-NB	306	<.001***

5 Conclusions

- All classifiers exhibit statistically significant performance differences compared to the majority classifier
- **K-nearest Neighbors** classifier stands out as the optimal choice based on its higher mean accuracy and low standard deviation across trials
- Non-linear classifiers are more suitable than linear classifiers

6 Limitations & Future Work

- **Database Shortage** -> implement stricter image selection criteria, conduct studies with larger and more diverse participant samples, and explore alternative annotation methods to enhance the quality, reliability, and generalizability of affective datasets
- **Limited Dimensions ->** expand the dimensions used for emotion classification and prediction, to advance the understanding of emotions, and improve the accuracy of emotion prediction models
- **ML Models ->** optimize hyperparameters, explore various machine learning algorithms such as SVM and more advanced techniques such as deep learning to improve accuracy and robustness in affective prediction tasks

References

[1]: C.J. Solomon and T.P. Breckon. Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab. Wiley-Blackwell, 2010.:
[2]: Kuan-Chuan Peng, Tsuhan Chen, Amir Sadovnik, and Andrew Gallagher. A mixed bag of emotions: Model, predict, and transfer emotion distributions. pages 860–868, 06 2015.
[3]: Terry M. Libkuman, Hajime Otani, Rosalie Kern, Steven G. Viger, and Nicole Novak. Multidimensional normative ratings for the international affective picture system. Behavior Research Methods, 39(2):326–334, May 2007.
[4]: Cannarile, Francesco & Compare, Michele & Baraldi, Piero & Diodati, G & Quaranta, Vincenzo & Zio, Enrico. (2019). Elastic Net Multinomial Logistic Regression for Fault Diagnostics of on-board Aeronautical Systems. Aerospace Science and Technology. 94. 10.1016/j.ast.2019.105392.