2026 International Conference on Automatic Face and Gesture Recognition (FG)

Facial Expression Features of Deception in Dynamic Naturalistic Social Interactions

Ailian Li[1,2], Jingting Li[1,2], Su-Jing Wang[1,2], Ye Liu[1,2,*]
[1] State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
[2] Department of Psychology, University of the Chinese Academy of Sciences, Beijing 101408, China

Abstract— **Facial expressions have been proposed as potential cues for deception detection in human interaction, yet their validity remains controversial. To address this debate, the present study systematically examined whether specific facial expression characteristics differ between deceptive and honest responses in a face-to-face interrogation competition paradigm. Three facial expression characteristics were analyzed: micro-expressions, facial asymmetry, and upper-lower facial dissociation. Facial expressions were annotated using the Facial Action Coding System (FACS), and machine learning analyses were subsequently conducted based on the action unit (AU) features. The results showed that micro-expressions were infrequent even when an extended temporal threshold was applied, and no significant difference in overall facial asymmetry was observed between deceptive and honest responses. Machine learning models trained based on AU features achieved moderate predictive performance, with RUSBoost demonstrating the highest accuracy in identifying deceptive responses. In general, facial cues provide limited yet informative signals for deception detection, and the findings offer interdisciplinary insights into the utilization of facial information in automated deception detection research.**

I. INTRODUCTION

Deception is a common phenomenon in social interaction and occurs in various communication contexts. Although some deceptive behaviors may be harmless, deception can also undermine social trust and lead to serious consequences in certain situations [27]. Therefore, accurately detecting deception is of considerable theoretical and practical importance. However, decades of research have consistently shown that human accuracy in deception detection is not significantly different from the chance level [2], [14]. The persistent difficulty in accurately detecting deception has prompted researchers to search for behavioral cues that may serve as valid indicators.

Among various nonverbal behaviors, facial expressions have been proposed as potentially informative cues for deception detection. Building on Darwin's early observations that certain facial muscles may escape voluntary control under intense emotion, Ekman formalized this idea in the inhibition hypothesis and further elaborated it in his leakage hypothesis [5]. According to these hypotheses, when individuals attempt to conceal their true emotions during deception, they may fail to fully suppress spontaneous facial muscle activation, resulting in a leakage of the underlying emotional state. Such leakage during deception may manifest in multiple ways. In the present study, we consider three potential manifestations: micro-expressions, facial asymmetry, and dissociation between upper and lower facial movements.

Despite their theoretical appeal, empirical findings regarding these facial cues for deception detection remain inconsistent. Micro-expressions, often defined as brief facial movements lasting less than 500 ms [28], are widely discussed as indicators of concealed emotion [6]. However, the observed incidence of micro-expressions is too low to permit meaningful statistical analysis of their role in deception detection [17]. Moreover, relatively few studies have directly examined whether micro-expressions can reliably discriminate between deceptive and truthful behavior. The ecological validity and practical utility of micro-expressions as reliable deception cues therefore remain debated [25].

In addition to micro-expressions, facial asymmetry and dissociation between upper and lower facial regions have also received attention. As a result of the underlying cognitive neural mechanisms, deliberately posed or regulated facial expressions tend to display heightened asymmetry [7], [20] and reduced coordination between upper and lower facial movements [5] relative to spontaneous expressions. However, empirical studies examining whether these features can serve as cues for deception detection remain relatively limited.

Despite its weak systematic empirical evidence in psychological research, the use of facial cues has become a prominent trend of automated deception detection research in computer science. While such data-driven approaches have achieved promising classification performance [13], [3], they often lack clear theoretical grounding regarding how specific facial structures relate to deceptive behavior, thereby limiting interpretability and raising concerns about robustness across contexts.

The present study addresses these gaps by examining facial cues in a face-to-face interrogation paradigm that allows participants to freely choose whether to respond honestly or deceptively. Using manual coding based on the Facial Action Coding System (FACS) [8], we systematically analyzed three facial expression characteristics including micro-expressions, facial asymmetry, and upper-lower facial dissociation under deceptive and honest conditions. In addition, we trained

This work was supported by the National Key Research and Development Program of China (2024YFE0202900); the National Natural Science Foundation of China (62276252, 62476269); the Youth Innovation Promotion Association, Chinese Academy of Sciences; and the Scientific Foundation of the State Key Laboratory of Cognitive Science and Mental Health (E5CX1201GZ).
* Corresponding author: liuye@psych.ac.cn

machine learning models based on action unit (AU)-level features to explore their predictive utility in distinguishing deception from honesty.

In summary, the contributions in this work include:

- Empirically evaluating the validity of facial expression characteristics (including micro-expressions, facial asymmetry, and upper-lower facial dissociation) for deception detection in a highly natural social interaction context.
- Training and comparing multiple machine learning models using AU features, to evaluate their predictive utility in distinguishing deceptive from honest responses.
- Integrating psychological theory and computational modeling, bridging the gap between theoretically grounded facial cues and data-driven classification approaches.

II. RELATED WORKS

This section reviews previous research on facial cues in deception detection, focusing on empirical findings regarding specific facial expression features, followed by research that applies computational methods to deception detection based on facial information.

A. *Micro-expressions and Deception*

Micro-expressions are typically defined as brief, low-intensity facial movements [10]. Because such facial actions are often difficult for an individual to fully suppress or voluntarily control [30], they may reveal genuine emotions when the individual attempt to conceal their true feelings [5]. For this reason, micro-expressions have been proposed as potential cues for detecting deception. Although this idea has received considerable attention in both scientific community and the media [10], systematic empirical research directly examining the relationship between micro-expressions and deception remains relatively limited [17], [25].

Among the existing studies, the research group led by Porter and ten Brinke conducted some of the earliest systematic investigations of the relationship between micro-expressions and deception. In an emotion-elicitation experiment, they found that micro-expressions appeared in only about 2% of the video segments [17]. In a subsequent study analyzing realistic high-stake interview situations, ten Brinke and Porter [24] identified only a very small number of micro-expression events (16 instances) across approximately 70,000 video frames, making it difficult to statistically distinguish liars from truth-tellers. These findings raised questions about the effectiveness of micro-expressions as stable cues for deception detection.

However, some researchers have suggested that these findings may be partly attributable to the operational definition of micro-expressions. In their studies, Porter and ten Brinke adopted the extremely strict temporal criterion (1/25-1/5 s) originally proposed by Ekman, which may have excluded many facial movements that were slightly longer in duration but still very brief. Matsumoto and Hwang (2018) compared the discriminative validity of micro-expressions under different temporal criteria ($\leq$200 ms to $\leq$500 ms) and found that when the operational definition was extended to approximately 400-500 ms, negative micro-expressions showed greater effectiveness in distinguishing liars from truth-tellers. This finding suggests that an extended temporal threshold (e.g., less than 500 ms) may better capture the potential role of micro-expressions in deception detection. In addition, some researchers have argued that the experimental context may also influence the likelihood of micro-expression occurrence. Most laboratory studies involve relatively low-stakes tasks in which participants lack strong emotional arousal, which may reduce the occurrence frequency of micro-expressions [23]. In contrast, high-stakes contexts may induce stronger emotional arousal and thus increase the likelihood of micro-expression occurrence. In such contexts, micro-expression features may be better able to distinguish between deceptive and honest behavior. For example, Shen et al. [19] analyzed video data from a high-stakes television game show, and found that the duration of fear-related micro-expressions was significantly shorter in deceptive clips than in truthful clips.

To date, existing research has yet to reach consistent conclusions regarding the role of micro-expressions in deception detection. On the one hand, some studies suggest that their low occurrence frequency limits their usefulness as valid cues for deception detection. On the other hand, other research indicates that under varying temporal definitions and situational conditions, micro-expressions may still provide meaningful information related to deceptive behavior.

B. *Faical Asymmetry, Upper-Lower Facial Dissociation, and Deception*

Left-right facial asymmetry has been suggested to be related to the authenticity of emotional expressions and may therefore serve as a potential cue for deception detection. Previous studies have found that when individuals deliberately produce expressions that are inconsistent with their genuine emotions, facial movements tend to show a greater degree of asymmetry between the left and right sides of the face [7], [20]. However, relatively few studies have directly examined facial asymmetry in deception contexts. Using the previously mentioned high-stakes deception dataset, Shen et al. (2021) [19] found that facial expressions were significantly more asymmetric in the lying condition than in the truth-telling condition.

In addition, upper-lower facial dissociation may also serve as a cue for detecting deception. Research suggests that different facial regions show a degree of functional specialization in emotional expression: the lower face is more frequently involved in the expression of positive emotions such as happiness, while the upper face more readily reflects emotions such as surprise or fear [18]. In situations that require emotional regulation, this functional differentiation may lead to inconsistent emotional signals across facial regions. For instance, in deceptive contexts, individuals may display socially appropriate smiles in the lower face while the upper face remains neutral or reveals traces of their genuine emotions. Experimental studies have also shown that, compared with liars, truth-tellers are more likely to display genuine smiles involving the co-activation of the zygomaticus major and orbicularis oculi muscles [9], [15].

Overall, both left-right facial asymmetry and dissociation between the upper and lower face may reveal inconsistencies between expressed and underlying emotions. However, the existing evidence remains relatively limited, particularly for systematic examinations of these facial characteristics in realistic deception contexts.

C. Facial Behavior and Automated Deception Detection

With the development of computer vision and machine learning, researchers have begun to develop automated deception detection tools based on visual information, extracting facial cues as features for machine learning models [21], [29].

In some studies, facial movements are manually annotated using coding systems such as MultiModal Interfaces (MUMIN) [1] and then used as input features for machine learning classifiers [16], [22]. Other studies rely on automated facial behavior analysis tools to extract AUs, which represent facial muscle activations and can capture subtle facial expressions and micro-expressions from video data [12], [19]. In addition, [23] derive visual features from facial landmarks and texture patterns to capture dynamic facial events such as eyebrow motion, mouth movement, and facial wrinkles.

Recent studies apply deep learning models, such as convolutional neural networks (CNNs) and 3D CNNs, to learn facial representations directly from video sequences and use these representations for deception detection [13], [3]. In addition, multimodal systems have been proposed to integrate facial features with other behavioral signals, such as audio and text, to improve the performance of deception detection [11].

Although some researches [13], [3] reported relatively high classification accuracy, model performance often relied heavily on specific datasets, experimental contexts, and feature selection strategies. This raises concerns about their generalizability. Furthermore, while existing studies have conducted fine-grained extraction and modeling of facial features, there remains a lack of systematic theoretical integration regarding the relationship between specific facial features and deception. Consequently, establishing a closer link between theory-driven facial coding frameworks and data-driven machine learning models is essential for developing more interpretable and theoretically informed deception detection systems.

III. METHODS

A. Participants

Thirty-eight undergraduate and graduate volunteers (32 females; mean age = 21.18 years, *SD* = 2.10 years) participated in this experiment. They were randomly paired into 19 dyads, and members of each dyad self-reported that they were not acquainted with each other prior to the experiment. All participants signed an informed consent form before the task and received monetary compensation upon completion of the experiment. The study adhered to the Declaration of Helsinki and was approved by the Institutional Review Board (IRB) of the authors' institution.

B. Paradigm

The experiment adopted a face-to-face interrogation competition paradigm [4]. Within each dyad, the two participants took turns playing the roles of informant and detective. The detective asked a series of questions, and the informant could choose to answer "yes" or "no" either truthfully or deceptively. This design allowed informants to freely decide whether to lie on a trial-by-trial basis, thereby eliciting spontaneous deceptive behavior.

The questions available to the detective were presented on a printed question sheet and comprised three types:

- Baseline questions, including two autobiographical questions based on the informant's true personal information, such as "Are you 21 years old this year?". When asked these questions, the informant was required to answer "yes" truthfully. The purpose of the baseline questions was to ensure that the informant remained attentive and engaged throughout the task; additionally, they provided a reference for the detective to compare the informant's reactions to baseline versus other questions, thereby facilitating judgment.
- Autobiographical information questions, consisting of 40 questions about personal life experiences, such as "Have you ever moved?".
- Personal preference questions, consisting of 40 questions about individual preferences, such as "Do you like reading novels?".

For both the autobiographical information questions and the personal preference questions, the informant was free to choose whether to answer truthfully or deceptively.

To motivate deceptive behavior in informants and detection performance in detectives, participants received a base payment along with performance-based bonuses. For each trial, when the informant provided a deceptive response, an additional bonus was awarded to the detective if the deception was detected, and to the informant if the deception was undetected.

C. Procedure

The experiment consisted of two phases. In the first phase, participants completed an online questionnaire based on their true autobiographical information and personal preferences. The questionnaire included all questions that the detective would later ask during the task and served as a reference to determine the veracity of the response. In the second phase, participants completed a face-to-face interrogation task at the laboratory.

The task consisted of a practice phase and a formal experimental phase. During the practice phase, the participants performed the same task as in the formal phase to ensure that they understood the requirements of the task and were familiar with the experimental procedure. The formal experimental phase consisted of two blocks, each containing 80 trials, corresponding to the 80 questions collected in the questionnaire. The order of questions was randomized for each participant. After completing the first block, participants within each dyad exchanged roles (informant and detective) and proceeded to the second block.

In each trial, the detective asked questions aloud following the order listed on the question sheet, judged whether the informant was responding deceptively, and recorded the judgment on the sheet. The informant's task was to provide a brief "yes" or "no" answer to each question. Except for the baseline questions, the informants were free to decide whether to respond truthfully or deceptively.

Upon completion of the experiment, participants received a base payment along with performance-based bonuses according to the rules described above. The bonus rules were explained in the instructions prior to the experiment to enhance informants' motivation to deceive and detectives' motivation to detect deception. After the experiment, the experimenter compared the informant's recorded responses in the video recordings with the answers provided in the pre-experiment questionnaire to determine whether each response was honest or deceptive.

D. Data Acquisition

During the experiment, each dyad sat face-to-face across a table. A video camera was mounted on a tripod positioned behind the detector, directly facing the informant, to record the informant's full face throughout the task. The videos were recorded at a resolution of 1200 $\times$ 1080 pixels at 25 frames per second and stored in MP4 format for subsequent FACS-based facial action analysis.

IV. DATA PROCESSING

A. Data Preprocessing

Facial expression video segments were obtained from 38 participants, comprising 476,281 video frames in total.

1) Behavioral labelling: The data collected from the questionnaire prior to the experiment and were used as a ground truth to determine the veracity of the responses. Participants' responses during the experiment were extracted from the recorded videos and compared with their questionnaire answers. For each trial, responses consistent with the questionnaire were coded as honest (0), and responses inconsistent were coded as deceptive (1).

2) Video coding and reliability assessment: Facial expression videos were manually coded based on the FACS. Each facial expression event was annotated with its onset frame, apex frame, and offset frame. For each event, coders additionally provided a subjective judgment of the expressed emotion (six basic emotions or "other"), identified the involved AUs, and coded the intensity of each AU on a standard ordinal scale (A-E levels). Coders also evaluated whether each AU was expressed symmetrically or asymmetrically across the left and right sides of the face.

TABLE I
COUNTS OF MICRO-EXPRESSIONS ACROSS EXPERIMENTAL CONDITIONS AND FACIAL REGIONS

Facial regions	Conditions		
	Honest	Deceptive	Total
Upper face	33	10	43
Lower face	12	22	34
Full face	10	3	13

All videos were coded by four certified FACS coders. To establish coding consistency, all four coders jointly coded the first four samples. After acceptable agreement (80%) was achieved, coders 1 and 2 independently coded samples 5-21, and coders 3 and 4 coded samples 22-38.

Following the FACS manual and previous research, inter-coder reliability was evaluated using the F_1 score, Matthews correlation coefficient (MCC), and the S score (free-marginal kappa). For each AU, presence versus absence was coded as a binary variable. Coding results from one coder were treated as the reference standard, and coding results from the other coder were treated analogously to the output of a classification model. Accordingly, performance metrics commonly used to evaluate machine learning classifiers were adopted to assess inter-coder agreement. For the first four jointly coded samples, inter-coder reliability was high ($F_1 = 0.841, MCC = 0.837, S = 0.826$). Discrepancies in the subsequent independently coded samples were resolved through discussion within each coder pair to reach consensus.

V. RESULTS

A. Deceptive Behavior

Participants produced an average of 24.29 deceptive responses across the task (*SD* = 13.50), with the number of deceptive responses ranging from 5 to 61. The mean deception rate was 31.1% (*SD* = 17.2%).

B. Facial Expression Results

The subsequent analyses focused on facial expression characteristics differentiating deceptive from truthful responses, specifically examining three types of facial features based on the analytical framework described above: micro-expressions, left-right facial asymmetry, and upper-lower facial dissociation.

1) Micro-expressions: Following previous research [28], micro-expressions were operationally defined as facial expressions lasting no more than 500 ms. Based on the FACS coding results, a total of 89 micro-expressions were identified across 2,984 trials, with a mean duration of 379 ms.

Given the relatively low occurrence frequency of micro-expressions, further inferential statistical analyses were not conducted. Instead, descriptive statistics were used to summarize their occurrence. The occurrence frequency of micro-expressions across experimental conditions and facial regions is reported in Table I, and detailed AU composition patterns are presented in Table II.

TABLE II
COUNTS OF MICRO-EXPRESSIONS BY SPECIFIC AU COMPOSITION PATTERNS ACROSS EXPERIMENTAL CONDITIONS AND FACIAL REGIONS

Facial regions	AU patterns	Conditions		
		Honest (n=2061)	Deceptive (n=923)	Total (n=2984)
Upper face	1+2	18	5	23
	2	5	1	6
	4	5	3	8
	5	1	1	2
	7	1	0	1
	1+2+5	1	0	1
	1	1	0	1
	2+6	1	0	1
	Total	33	10	43
Lower face	19+25+26	4	2	6
	39	4	0	4
	38	2	1	3
	14+17	2	0	2
	14+24	2	0	2
	11	1	0	1
	12	0	1	1
	12+25	0	1	1
	12+25+26	0	1	1
	14	0	1	1
	14+17+26+35	0	1	1
	14+21+24+31	1	0	1
	14+25+26+37	0	1	1
	16+25	0	1	1
	17	1	0	1
	17+18	1	0	1
	17+19+26	0	1	1
	18+23+26	1	0	1
	21	0	1	1
	21+31	1	0	1
	29	1	0	1
	31	1	0	1
	Total	12	22	34
Full face	1+2+25+26	2	0	2
	1+2+12	0	1	1
	1+2+23	1	0	1
	1+2+12+25+26	1	0	1
	1+2+7+12+23	1	0	1
	4+9	1	0	1
	4+6+12	1	0	1
	4+10+12	1	0	1
	4+10+25	0	1	1
	4+7+10+12	1	0	1
	6+7+10+17	0	1	1
	7+12+19+25+26	1	0	1
	Total	10	3	13
Total		63	26	89

Among upper-face micro-expressions, the most frequently observed AU combination was AU1+AU2. The second most frequently observed micro-expression involved AU4.

2) Facial asymmetry: Facial asymmetry was operationally defined as any facial expression event that contained at least one asymmetric AU, as indicated by the FACS coding.

To examine whether the proportion of asymmetric facial expressions was higher under deception than under honest responding at the individual level, analyses were conducted based on the AU-level asymmetry annotations for each facial expression event. For each participant, the number of expression events containing asymmetric AUs was calculated separately for the deception and honest conditions and divided by the total number of trials in the corresponding condition, yielding the proportion of asymmetric expressions for each condition.

Descriptive statistics showed that the mean proportion of asymmetric expressions was 0.096 (SD = 0.116) in deceptive trials and 0.106 (SD = 0.114) in honest trials. A Wilcoxon signed-rank test was conducted to compare within-subject differences in the proportion of asymmetric expressions between the two conditions. The results indicated no significant difference between deceptive and honest conditions ($V = 222$, $p = 0.438$). The results are illustrated in Fig. 1.

3) Upper-lower facial dissociation: To explore potential patterns of upper-lower facial dissociation in deceptive ex-

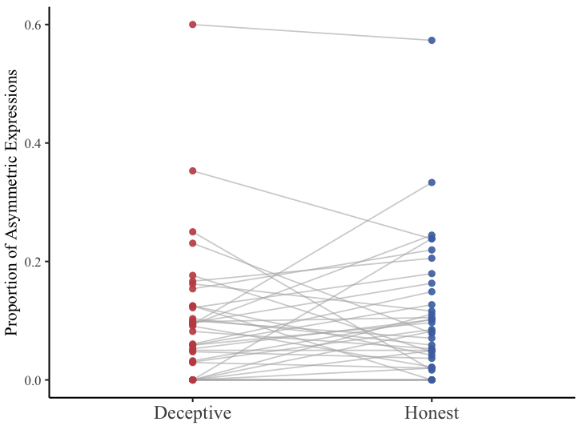

Fig. 1. Proportion of Asymmetric Expressions in Deceptive and Honest Conditions.

TABLE III
MOST FREQUENT AU COMPOSITION PATTERNS OF MACRO-EXPRESSIONS ACROSS CONDITIONS AND FACIAL REGIONS

Facial regions	AU patterns	Conditions		
		Honest (n=2061)	Deceptive (n=923)	Total (n=2984)
Upper face	1+2	54	19	73
	2	26	8	34
	4	9	14	23
	1+2+5	7	8	15
	1	10	3	13
	7	8	4	12
	...	...	...	...
	Total	132	63	195
Lower face	14	105	35	140
	14+24	78	21	99
	17	28	6	34
	12	23	11	34
	14+17+24	28	2	30
	12+25	19	9	28
	...	...	...	...
	Total	678	228	906
Full face	6+7+12+25	27	14	41
	7+12+25	32	4	36
	6+7+12+25+26	20	11	31
	6+7+10+12+25	3	19	22
	...	...	...	...
	Total	343	169	512
Total		1153	460	1613

Note: Due to the large number of possible AU patterns, only the most frequently observed patterns in each condition are reported.

pressions, we conducted an exploratory analysis of AU composition patterns. As the AU patterns of micro-expressions have been presented in Table II , Table III presents the most frequently observed AU patterns for the other facial expressions (macro-expressions) within each facial region and condition.

C. Machine Learning Results

1) Configuration: Classification models were trained to distinguish deceptive from honest responses using facial AU-based features derived from manually coded FACS data. For each trial, AU annotations were available at the level of expression events and aggregated into trial-level features.

TABLE IV
PERFORMANCE OF MACHINE LEARNING CLASSIFIERS FOR DISTINGUISHING DECEPTIVE AND HONEST RESPONSES

Algorithm	CV.ACC	ACC	Prec.	Rec.	Spec.	F_1	BA
SVM	0.690	0.813	0.953	0.417	0.991	0.580	0.704
RF	0.702	0.711	0.940	0.068	0.998	0.127	0.533
AdaBoost	0.702	0.730	0.860	0.153	0.989	0.259	0.571
RUSBoost	**0.614**	**0.702**	**0.515**	**0.620**	**0.739**	**0.562**	**0.679**

Note. CV.ACC = mean accuracy across five-fold cross-validation; ACC = accuracy on the full dataset; Prec. = precision; Rec. = recall; Spec. = specificity; F_1 = F_1-score; BA = balanced accuracy. SVM = support vector machine; RF = random forest; AdaBoost = adaptive boosting; RUSBoost = random undersampling boosting.

For each AU, we extracted 5 features, including maximum intensity, mean intensity, weighted mean intensity, asymmetry, and binary presence indicators. Among these, the mean intensity was calculated as the average AU intensity across all expression events within a trial, while the weighted mean intensity was calculated by weighting the intensity of each event by the duration of that event within the trial. A total of 38 AUs were included, resulting in an initial feature space of 190 dimensions. In addition to these AU-based features, metadata (SubjectID, question number, and trial number) were also included. Features with zero variance across trials were excluded, resulting a final feature set of 175 dimensions. All features were normalized before model training.

In total, the dataset consisted of 2,984 trials, including 2,061 honest responses and 923 deceptive responses. The task was formulated as a binary classification task (honest vs. deceptive), and model performance was evaluated using 5-fold cross-validation. Four classifiers were evaluated: support vector machine (SVM), random forest (RF), AdaBoost, and RUSBoost.

2) Classification Reuslt: Performance metrics for all models are reported in Table IV. Although RUSBoost did not achieve the highest overall cross-validated accuracy, it showed the best performance in terms of recall for deceptive trials. Given that the primary goal of this analysis was to identify deceptive responses, recall for the deception class was considered the most relevant performance metric.

For the RUSBoost model, the overall training accuracy was 70.2%, and the mean accuracy across five-fold cross-validation was 61.4%. The model achieved a recall of 61.97% and a specificity of 73.9% The F_1 score was 56.2%, and the balanced accuracy was 67.9%. As shown in Fig. 2, the confusion matrix indicated that the model still exhibited a non-negligible rate of false negatives for deceptive trials, whereas performance for honest trials was relatively better.

Feature importance analysis indicated that model predictions were primarily driven by the top 20% of features. Fig. 3 shows the 10 most important features among them. These features fell into two broad categories: (1) subject- and trial-level information (e.g., Subject ID and question number), and (2) AU-related features, including AU14 asymmetry, AU10 weighted mean intensity, AU10 mean intensity, AU6 mean

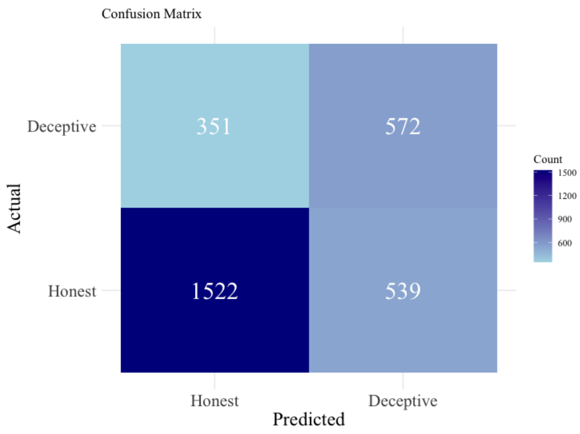

Fig. 2. Confusion matrix for the RUSBoost model.

intensity, AU2 asymmetry and AU12 asymmetry.

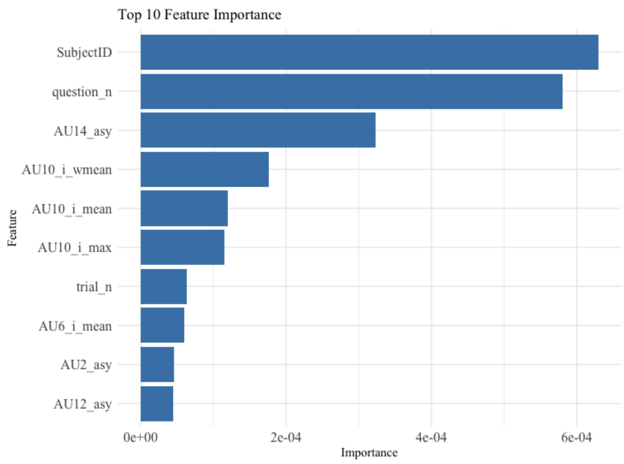

Fig. 3. Top 10 most important features in the classification model.

VI. DISCUSSION

The present study examined facial expression characteristics in a face-to-face interrogation competition paradigm. Overall, the findings suggest that facial behavioral cues provide limited but informative signals for distinguishing deceptive from honest responses.

Micro-expressions have been widely discussed as potential indicators of concealed emotion [10]. However, even with an extended temporal threshold ($\leq$500 ms), micro-expressions were still observed infrequently in the present study, and instances mainly involved upper-face AUs such as AU1 + AU2 and AU4. Due to the limited number of micro-expressions, no statistical comparison between deceptive and honest conditions was conducted. This low occurrence frequency may reflect the relatively low-risk nature of the task, in which participants may not have experienced sufficiently strong emotional arousal or conflict to elicit micro-expressions. Consistent with theories suggesting that micro-expressions arise from the conflict between emotional arousal and inhibitory control [5], many prior studies have relied on paradigms that induce strong emotions while requiring participants to suppress expressions [26], [28]. The low occurrence frequency observed here highlights the potential limitations of micro-expressions as reliable cues for deception detection in naturalistic, low- to medium-risk settings.

The present study did not find a significant difference in the overall proportion of asymmetric facial expressions between deceptive and honest responses. One possible explanation concerns difference in operationalization. In this study, asymmetry was defined at the level of expression events based on whether an expression contained at least one asymmetric AU according to FACS coding. In contrast, some previous studies have used more fine-grained dynamic measures of left-right coordination. For example, Shen et al. [19] quantified synchronization between left and right facial movements using wavelet coherence between corresponding facial landmarks. Such approaches capture continuous temporal coordination between the two sides of the face rather than the presence of asymmetric AUs within discrete expression events. Interestingly, the machine learning analyses identified several asymmetric AU features, such as asymmetric activations of AU12 and AU14, as relatively important predictors of deception. This suggests that the asymmetry of specific AUs may be more informative for deception detection than the overall proportion of asymmetric expressions.

The present study also explored upper-lower facial activation patterns. However, a clear and quantifiable operational definition of upper-lower facial dissociation is still lacking in the current literature, which also limits the interpretability and replicability of the present findings. Future research may consider at least two potential approaches. One possibility is to define dissociation based on the emotional mappings of AUs. If upper- and lower- face AUs within the expression are typically associated with different emotional states, the expression may be considered upper-lower dissociated. However, this approach is limited by the many-to-many relationship between AUs and emotions, as AUs do not map uniquely onto discrete affective categories. Alternatively, trained coders could manually assess the degree of upper-lower facial dissociation in each expression. While this approach provides a direct measure of dissociation, it is susceptible to the coder's subjectivity and can be time-consuming. Therefore, developing a reliable and theoretically grounded operational definition of upper-lower facial dissociation remains an important direction for future work.

The machine learning results suggest that AU-based facial features contain some predictive information for distinguishing deceptive from honest responses. Several classifiers achieved accuracy ranging from approximately 0.70 to 0.81, which is broadly comparable to results reported in previous computational studies of deception detection using facial or multimodal behavioral cues (e.g., [16], [12], [23], [14]). However, current machine learning analysis primarily utilize individual AUs as features, without considering coordinated patterns of multiple AUs. Given that facial expressions often emerge as integrated configurations, future research may further examine specific AU combinations to improve the interpretability of the model. In addition, performance metrics revealed a clear imbalance, classifiers showed high specificity for honest responses but relatively low recall for deceptive responses. Among the models, RUSBoost produced the highest recall for deception, suggesting that algorithms designed to address class imbalance may be more suitable for deception detection tasks. Notably, Subject ID emerged as the most important feature, likely reflecting individual differences rather than generalizable cues for deception. This highlights the need for larger and more diverse samples to ensure models capture deception patterns rather than individual traits. Overall, these findings align with previous research indicating that accurately identifying deception remains challenging when relying on facial signals alone [25].

Several limitations should be acknowledged. First, the present experimental paradigm represents a relatively low-risk context and may not have sufficiently induced strong emotional conflict, which may have resulted in fewer observable facial expression cues. Therefore, future studies should further examine the effectiveness of facial behavioral cues for deception detection across different contexts. Second, individual differences in facial behavior may have influenced the results. Some participants tended to exhibit certain facial movement patterns regardless of experimental condition. Such individual-level behavioral tendencies may partially attenuate differences between conditions. Future research may consider increasing sample size to better capture population-level patterns. In addition, the current sample consisted primarily of young adults, which limits the generalizability of the findings to other age groups. Future research should investigate whether facial cues to deception vary across different age ranges. Furthermore, due to the relatively small sample size, it was not possible to examine potential gender differences. And this issue should be further examined in future research with a larger sample size. Finally, deception is a complex behavioral process. Relying solely on facial features may be insufficient to achieve stable and generalizable detection performance. As suggested by studies in computer science, multimodal information fusion—such as combining audio, text, and other behavioral signals—may improve the accuracy and robustness of deception detection.

VII. Conclusions

This study investigated facial expression features extracted via manual FACS coding as predictors for deception detection in a face-to-face interaction paradigm. The results indicate that micro-expressions, facial asymmetry, and upper-lower facial dissociation patterns provide limited yet informative cues for distinguishing deceptive from honest responses. Machine learning classifiers trained on these features achieved moderate predictive performance, suggesting that facial behavioral signals carry detectable information, although accurately detecting deception remains challenging when relying solely on facial cues. Overall, the findings advance deception detection research by integrating theory-driven facial feature extraction with data-driven classification. Given the complexity and context dependence of deceptive behavior, future studies may benefit from combining structured behavioral analysis with multimodal cues.

References

[1] J. Allwood, L. Cerrato, K. Jokinen, C. Navarretta, and P. Paggio. The mumin coding scheme for the annotation of feedback, turn management and sequencing phenomena. *Language Resources and Evaluation*, 41(3):273–287, 2007.

[2] C. F. Bond and B. M. DePaulo. Accuracy of deception judgments. *Personality and Social Psychology Review*, 10(3):214–234, 2006.

[3] M. Ding, A. Zhao, Z. Lu, T. Xiang, and J.-R. Wen. Face-focused cross-stream network for deception detection in videos. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7794–7803, 2019.

[4] Z. Dong, G. Wang, S. Lu, L. Dai, S. Huang, and Y. Liu. Intentional-deception detection based on facial muscle movements in an interactive social context. *Pattern Recognition Letters*, 164:30–39, 2022.

[5] P. Ekman. Darwin, deception, and facial expression. *Annals of the New York Academy of Sciences*, 1000(1):205–221, 2003.

[6] P. Ekman. *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*. W. W. Norton and Company, 2009.

[7] P. Ekman, W. V. Friesen, and J. C. Hager. The symmetry of emotional and deliberate facial actions. *Psychophysiology*, 18(2):101–106, 1981.

[8] P. Ekman, W. V. Friesen, and J. C. Hager. *Facial Action Coding System*. A Human Face, Salt Lake City, UT, 2002.

[9] P. Ekman, W. V. Friesen, and M. O'Sullivan. Smiles when lying. *Journal of Personality and Social Psychology*, 54(3):414–420, 1988.

[10] P. Ekman and M. O'Sullivan. From flawed self-assessment to blatant whoppers: The utility of voluntary and involuntary behavior in detecting deception. *Behavioral Sciences & the Law*, 24(5):673–686, 2006.

[11] D. Grabowski, K. Łuczaj, and K. Saeed. Multimodal behavioral sensors for lie detection: Integrating visual, auditory, and generative reasoning cues. *Sensors*, 25(19):6086, 2025.

[12] M. Jaiswal, S. Tabibu, and R. Bajpai. The truth and nothing but the truth: Multimodal analysis for deception detection. In *IEEE International Conference on Data Mining Workshops*, pages 938–943, 2016.

[13] G. Krishnamurthy, N. Majumder, S. Poria, and E. Cambria. A deep learning approach for multimodal deception detection, 2018. arXiv:1803.00344.

[14] M. Monaro, S. Maldera, C. Scarpazza, G. Sartori, and N. Navarin. Detecting deception through facial expressions in a dataset of video-taped interviews: A comparison between human judges and machine learning models. *Computers in Human Behavior*, 127:107063, 2022.

[15] A. Mondal, P. Mukhopadhyay, N. Basu, S. Bandyopadhyay, and T. Chatterjee. Quantitative analysis of euclidean distance to complement qualitative analysis of facial expression during deception. *Industrial Psychiatry Journal*, 25(1):78, 2016.

[16] V. Pérez-Rosas, M. Abouelenien, R. Mihalcea, and M. Burzo. Deception detection using real-life trial data. In *Proceedings of the ACM International Conference on Multimodal Interaction*, pages 59–66, 2015.

[17] S. Porter and L. ten Brinke. Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions. *Psychological Science*, 19(5):508–514, 2008.

[18] E. D. Ross, C. I. Prodan, and M. Monnot. Human facial expressions are organized functionally across the upper-lower facial axis. *The Neuroscientist*, 13(5):433–446, 2007.

[19] X. Shen, G. Fan, C. Niu, and Z. Chen. Catching a liar through facial expression of fear. *Frontiers in Psychology*, 12:675097, 2021.

[20] M. Skinner and B. Mullen. Facial asymmetry in emotional expression: A meta-analysis of research. *British Journal of Social Psychology*, 30(2):113–124, 1991.

[21] T. E. Slowe and V. Govindaraju. Automatic deceit indication through reliable facial expressions. In *IEEE Workshop on Automatic Identification Advanced Technologies*, pages 87–92, 2007.

[22] F. Soldner, V. Pérez-Rosas, and R. Mihalcea. Box of lies: Multimodal deception detection in dialogues. In J. Burstein, C. Doran, and T. Solorio, editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 1768–1777. Association for Computational Linguistics, 2019.

[23] L. Su and M. Levine. Does "lie to me" lie to you? an evaluation of facial clues to high-stakes deception. *Computer Vision and Image Understanding*, 147:52–68, 2016.

[24] L. ten Brinke and S. Porter. Cry me a river: Identifying the behavioral consequences of extremely high-stakes interpersonal deception. *Law and Human Behavior*, 36(6):469–477, 2012.

[25] A. Vrij, M. Hartwig, and P. A. Granhag. Reading lies: Nonverbal communication and deception. *Annual Review of Psychology*, 70:295–317, 2019.

[26] S. Wang et al. Micro-expression recognition based on eeg signals. *Biomedical Signal Processing and Control*, 86:105037, 2023.

[27] Z. Wu, B. Singh, L. S. Davis, and V. S. Subrahmanian. Deception detection in videos. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

[28] W.-J. Yan, Q. Wu, J. Liang, Y.-H. Chen, and X. Fu. How fast are the leaked facial expressions: The duration of micro-expressions. *Journal of Nonverbal Behavior*, 37(4):217–230, 2013.

[29] Z. Zhang, V. Singh, T. E. Slowe, S. Tulyakov, and V. Govindaraju. Real-time automatic deceit detection from involuntary facial expressions. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–6, 2007.

[30] G. Zhao, X. Li, Y. Li, and M. Pietikäinen. Facial micro-expressions: An overview. *Proceedings of the IEEE*, 111(10):1215–1235, 2023.