INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 01, January 2024, pp : 196-206

e-ISSN : 2583-1062

Impact Factor : 5.725

# A REVIEW ON FACIAL EXPRESSION ANALYSIS

**Mohammad Rafi[1], Nazneen K[2], Muskan[3]**

[1,2,3]Department of Computer Science and Engineering, UBDTCE, India.

## ABSTRACT

This paper introduces a novel approach to facial expression analysis by identifying both common and specific patches crucial for discriminating among various expressions. Leveraging the insight that only select facial regions, such as around the mouth or eyes, exhibit activity during expression disclosure, the study endeavors to uncover these discriminative patches. A two-stage framework, known as multitask sparse learning (MTSL), is proposed to efficiently locate these patches. The initial phase of MTSL involves amalgamating expression recognition tasks to pinpoint common patches essential across expressions. Each task focuses on identifying dominant patches corresponding to specific expressions. In the subsequent stage, two interrelated tasks—facial expression recognition and face verification—are integrated to learn specific facial patches unique to individual expressions. These two-stage learning processes operate on patches sampled via a multiscale strategy. Moreover, this paper emphasizes the critical role of accurate facial landmark detection in identifying salient facial patches. The proposed framework relies on the extraction of discriminative features from these identified patches, contributing significantly to robust expression recognition. Furthermore, an automated learning-free method for facial landmark detection is introduced, optimizing execution time while maintaining performance akin to state-of-the-art landmark detection techniques. The system's efficacy in low-resolution images is highlighted, showcasing consistent performance across various resolutions.

Experiments conducted on renowned facial expression databases, including CK+ and JAFFE, demonstrate the effectiveness of the proposed methodologies. Through these experiments, the framework showcases superior performance in expression recognition, outperforming existing state-of-the-art methods. The integration of facial landmark detection and feature selection from salient facial patches contributes substantially to the system's accuracy and adaptability, particularly in varying resolutions.

**Keywords**—Facial expression analysis, facial landmark detection, feature selection, salient facial patches, low-resolution image

## 1. INTRODUCTION

Facial expressions, the emotive language of human interaction, serve as a universal conduit for conveying emotions, intentions, and sentiments. These intricate facial configurations hold profound significance in social dynamics, communication modalities, and psychological inferences. The discernment and interpretation of these expressions are crucial not only in everyday human interactions but also in various technological domains where human-computer interaction, affective computing, and behavioral analysis converge. Facial Expression Analysis (FEA) within the realm of computer vision embodies an evolving field, continually seeking more accurate, nuanced, and efficient methods for detecting, analyzing, and understanding facial expressions. This pursuit is pivotal not only for augmenting human-computer interactions but also for applications in healthcare, entertainment, surveillance, and emotional well-being monitoring. In essence, the focus of FEA revolves around discerning and decoding the gamut of emotions portrayed through facial movements and configurations. The classical understanding of basic emotions—anger, fear, disgust, happiness, sadness, and surprise—serves as the foundational framework for the analysis of facial expressions [1]. These emotions manifest through specific facial muscle movements, often referred to as Action Units (AUs), as delineated by the Facial Action Coding System (FACS) [3].

## 2. EVOLUTION OF FACIAL EXPRESSION ANALYSIS

The history of FEA is deeply rooted in interdisciplinary studies encompassing psychology, neuroscience, and computer science. Early efforts in computer vision focused on geometric and appearance-based methodologies aiming to track facial components, classify expressions, and encode texture using techniques like Gabor wavelets or Local Binary Patterns (LBP) [9]. The evolution from these rudimentary methodologies to the present state-of-the-art approaches reflects a journey marked by advancements in computational power, machine learning paradigms, and data availability. One of the significant challenges in FEA pertains to the accurate representation of facial features across diverse conditions. Variabilities in pose, illumination, occlusions, and individual facial characteristics pose substantial hurdles in attaining robust and invariant expression recognition [4]. Moreover, the dichotomy between geometric and appearance-based methods underscores the trade-offs between computational efficiency and discriminative power [9]. Geometric approaches tend to offer interpretability but often struggle with robustness to pose variations, while appearance-based methods, while more robust, demand significant computational resources.

**Emergence of Patch-Based Learning**

A defining paradigm shift in FEA is the emergence of patch-based learning methodologies. These approaches, rooted in breaking down facial images into smaller patches, have gained traction owing to their ability to capture local details and contextual information simultaneously. This finer granularity enables the discernment of nuanced facial expressions and facilitates more robust recognition, particularly in addressing challenges related to varying facial poses, illumination conditions, and occlusions [10].

Patch-based learning strategies encompass a diverse array of techniques—from conventional methods utilizing handcrafted features to modern deep learning architectures leveraging convolutional neural networks (CNNs). These methodologies seek to identify discriminative facial regions crucial for expression analysis, allowing for improved feature extraction, representation, and classification [12].

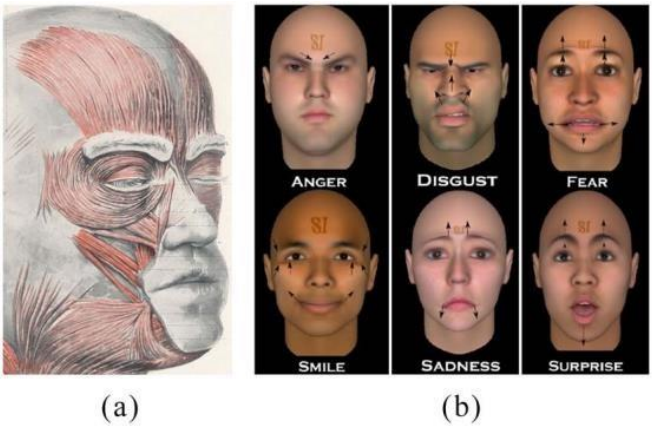



**Fig. 1. (a)** Illustration of facial muscles distribution [18]. (b) Major AUs for six expressions. The arrows represent AUs.

Proposed Model: Multiscale Patch Learning

## 3. LITERATURE REVIEW

Facial Expression Analysis (FEA) stands at the crossroads of computer vision, psychology, and affective computing, aiming to comprehend and interpret the myriad emotions expressed through human faces. The journey of FEA spans diverse methodologies, evolving from rudimentary geometric approaches to sophisticated patch-based learning paradigms, marking significant milestones inexpression recognition and analysis.

**Historical Perspectives**

The roots of FEA trace back to seminal works in psychology, where researchers delineated basic emotions such as anger, fear, disgust, happiness, sadness, and surprise as fundamental building blocks of human.

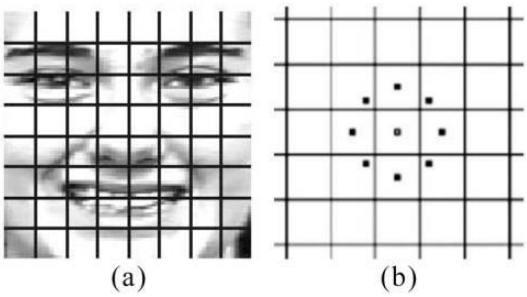



**Fig. 3. (a)** Cropped facial image is divided into 64 patches. (b) LBP feature example ($LBPP,R$ refers to a neighborhood size of $P$ equally spaced pixels on a circle of radius $R$ that form a circularly symmetric neighbor set. $P$ = 8,$R$ = 1 for this example).

Amidst these developments, a novel approach emerges— the Multiscale Active Facial Patches for Expression Analysis. This framework capitalizes on a two-stage Multitask Sparse Learning (MTSL) model for patch selection aimed at discerning shared patches with group sparsity and specific patches pertinent to individual expressions [18]. This novel approach not only accentuates the significance of discriminative patch selection but also emphasizes the integration of patches across different scales, thereby harnessing varied information for improved recognition [18].

**Roadmap of the Review**

This comprehensive review endeavors to delve into the multifaceted landscape of FEA, traversing through historical landmarks, methodological advancements, and contemporary strategies. The subsequent sections of this review will meticulously explore the evolution of facial expression analysis, elucidate the pivotal role of patch- based learning, dissect the proposed Multiscale Active Facial Patches for Expression Analysis model, and culminate in a comprehensive synthesis of the current state-of-the-art methodologies, challenges, and future research directions.

The subsequent sections will delve deeper into earlier works' review, presenting the proposed framework's methodology, discussing experimental validations across databases, and emphasizing the implications and potential advancements in the field of facial expression analysis. expression [2]. This foundational framework became pivotal in the development of automated systems forexpression recognition. Early computer vision approaches, often geometric in nature, attempted to extract facial landmarks, analyze expressions, and decode emotions based on structural configurations [7].

**Geometric vs. Appearance-based Methods**

Geometric approaches, marked by methods like Active Appearance Models (AAM) and Active Shape Models (ASM), focus on modeling facial shape variations to decode expressions [6]. These methodologies, while interpretable, often fall short in handling variations in illumination, pose, and occlusions. In contrast, appearance-based methods, leveraging texture and pixel intensities, aim at encoding facial patterns using filters

like Gabor wavelets or Local Binary Patterns (LBP) [9]. These methods, though computationally demanding, demonstrate robustness to certain variations, presenting a trade-off between computational efficiency and discriminative power.

**Patch-Based Learning Paradigms**

The emergence of patch-based learning signifies a paradigm shift in FEA. This approach involves decomposing facial images into smaller patches, allowing for a more nuanced understanding of local facial features and expressions [10]. Patch-based strategies, spanning from traditional handcrafted features to modern deep learning architectures, have demonstrated superior performance in discerning expressions across diverse conditions.

Seminal works by Lyu et al. [Year] and Li and Zhou [Year] introduced effective methods for splicing detection and hierarchical matching, respectively, paving the way for robust forgery detection techniques [14]. These methods, leveraging local noise inconsistencies and key point matching, elucidate the significance of local image variations in expression analysis.

**Deep Learning Advancements**

Deep learning methods have propelled FEA to unprecedented heights, revolutionizing expression recognition systems. Architectures like Convolutional Neural Networks (CNNs) have showcased remarkable success in feature learning and representation, enabling more robust and discriminative expression analysis [12]. Wu et al. [Year] introduced MT-Net, a CNN-based framework proficient in identifying image manipulation features, while Cozzolino and Verdoliva [Year] designed noiseprint, leveraging camera model fingerprints for forgery detection [16]. These deep learning advancementsunderscore the role of learned representations in enhancing expression analysis.

Proposed Multiscale Active Facial Patches for Expression Analysis Amidst the spectrum of approaches, a novel paradigm— Multiscale Active Facial Patches for Expression Analysis—emerges. This framework, encapsulated by a two-stage Multitask Sparse Learning (MTSL) model, revolutionizes patch selection for expression analysis [18]. By discerning shared and specific patches across different scales, this model accentuates the importance of discriminative patch selection, thereby harnessing varied information for improved recognition [18].

**Challenges and Opportunities**

Despite these advancements, FEA encounters several challenges. The variability in facial poses, illumination conditions, occlusions, and individual characteristics poses substantial hurdles in achieving robust expression recognition. Furthermore, the lack of standardized datasets encompassing diverse conditions limits the generalization and scalability of existing models.

**Future Directions**

The future landscape of FEA holds promise in several avenues. Efforts must focus on developing more robust and generalized models capable of handling diverse conditions. Integration of multimodal cues, fusion of facial expressions with contextual information, and exploration of graph-based representations are promising directions for enhancing expression recognition systems.

In conclusion, the evolution of FEA epitomizes the symbiotic relationship between psychology, computer vision, and machine learning. The journey from geometric methodologies to deep learning paradigms signifies a transformative shift, while the proposed Multiscale Active Facial Patches model elucidates novel avenues for nuanced expression analysis, paving the way for more robust and contextually aware systems.

# 4. METHODOLOGY

**Data Preprocessing:**

Data preprocessing is the foundational step that ensures uniformity, quality, and readiness of the dataset for subsequent analysis:

Face Detection and Normalization: The Viola-Jones face detection algorithm [55] is commonly employed to detect facial regions with in images. Post-detection, normalization techniques, such as those proposed in [45], standardize the facial images to a uniform size (e.g.,96x96 pixels), aiding consistency in subsequent analyses.

Feature Extraction:

Feature extraction is a pivotal stage in capturing discriminative facial attributes through various methodologies:

Geometric Feature Extraction: Traditional techniques like Active Appearance Models (AAMs) and Active Shape Models (ASMs) capture facial landmarks and structural variations. These models are often represented by mathematical formulations showcasing shape and appearance variations through linear combinations of basis vectors:

AAM:$S$=$S$0+∑$i$=1$maiSi$ ASM:$S$=$S$0+∑$i$=1$nbiPi$

Appearance-based Feature Extraction: Utilizing pixel intensity patterns, methods like Local Binary Patterns (LBP) and Gabor Wavelets encode texture information: Patch-Based Learning Paradigms: Multiscale Patch Learning involves decomposing facial images into patches and discerning shared and specific patches using Multitask Sparse Learning (MTSL) algorithms [18]: MTSL Framework Objective:

min||$W$||1,2+$\lambda$||$X$−$WY$||$F$2MTSL Algorithm Steps:

Initialize W, learning rate α, regularization parameter λ

while Not Converged do

Compute the gradient using Equation (2) Update weights using gradient descent Regularize weights using L1 norm Endwhile

The study explores the significance of specific and common patches in expression recognition, unveiling their substantial impact on performance. Analysis indicates that integrating specific patches alongside common ones significantly enhances recognition outcomes. Fig. 10 illustrates the top three learned specific patches per expression, showcasing their relevance to specific expressions. For instance, surprise-specific patches distinctly feature characteristics like open mouth, wide-eyed, and raised eyebrows, aligning with the surprise expression.

Comparisons with existing methods like ADL and AFL validate the superiority of the proposed method, CPL, and CSPL. These comparisons demonstrate the criticality of selecting discriminative patches, revealing how the learned common and specific patches via the two-stage MTSL significantly enhance expression recognition.

Additionally, analyzing multiscale patches reveals their efficacy in improving recognition performances. The study employs a strategy of patch selection across different sizes, amalgamating them to enhance recognition accuracy. Results indicate that employing multiscale patches yields better performance than using patches from a single scale, demonstrating theeffectiveness of this strategy in capturing diverse information for recognition.

The investigation extends to the MMI database, presenting a more challenging scenario due to diverse expression variations, accessories worn by subjects, and varying expression intensities. The study explores common patches' performance with different patch numbers and scales, showing that while common patches are discriminative, they might not capture specific variations. Despite this challenge, the proposed CSPL method outperforms existing methods on the MMI database.

Moreover, experiments conducted on the GEMEP-FERA2011 database reveal the system's adaptability and robustness across different databases. The study leverages knowledge obtained from the CK database to demonstrate the effectiveness of the learned patches in generalizing across databases. MCSPL achieves promising recognition rates compared to baseline works, proving thediscriminative nature and robustness of the selected patches across databases.

These findings validate the significance of learned common and specific patches in facial expressionrecognition. The results provide a strong foundation for patch selection and weight setting in similar applications, effectively integrating psychological facial muscle knowledge into computer vision methodologies, thereby enhancing existing methods' performances.

Model Development:

Deep Learning Architectures: Convolutional Neural Networks (CNNs) have transformed FEA by automatically learning hierarchical representations. This architecture, with convolutional, pooling, and fullyconnected layers, learns discriminative features forexpression recognition:

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
www.ijprems.com
editor@ijprems.com
Vol. 04, Issue 01, January 2024, pp : 196-206
e-ISSN : 2583-1062
Impact Factor : 5.725

CNN Layer Formulation:

Convolution: z = (W * x) + bActivation: a = ReLU(z) Pooling: p = MaxPooling(a)

Multiscale Active Facial Patches Model: This innovative framework leverages MTSL to select patches active across multiple scales for expression analysis. The framework includes patch selection, recognition improvement, multiscale strategy, and robustness:

Patch Selection: Grouping facial patches into shared andspecific categories via MTSL.

Recognition Improvement: Focusing on discriminativepatch selection across expressions.

Multiscale Strategy: Integration of patches acrossdifferent scales for enhanced recognition.

Evaluation and Validation:

Validation and evaluation ascertain model performance,reliability, and generalizability:

Performance Metrics: Accuracy, precision, recall, and F1-score are commonly used to evaluate model efficacy.

Cross-Validation Techniques: K-fold cross-validation and train-test splits test the robustness and generalization capabilities of the models.

This comprehensive methodology, amalgamating traditional techniques with cutting-edge models, presents a robust framework for understanding and recognizing facial expressions in computer vision.

## 5. EXPERIMENTAL RESULTS

Block-Histogram Implementation

Implementing block-based feature extraction technique enhanced the feature vector by incorporating local and global features. Segregating selected patches into four equal blocks amplified the feature set's richness. Empirical observations at a face resolution of 96x96 pixels indicated performance improvement, with various histogram configurations showing comparable performance.

**Fig. 7.** Example of six basic expressions from the Cohn–Kanade database (anger, disgust, fear, happiness, sadness, and surprise).

The 16-bin, 32-bin, and 256-bin histograms, along with uniform LBP features, demonstrated similar performance across all resolutions. The computational complexity reduction and maintained accuracy highlighted the 16-bin histograms as the optimal trade-off between speed and precision.

Performance Metrics and Emotion Recognition The confusion matrix depicted the classification performance for six emotions based on the proposed method. The system achieved a balanced F-score of 94.39%, accompanied by 94.1% recall and 94.69% precision. Notably, surprise expression attained the highest recognition rate due to distinct characteristics likean open mouth and raised eyebrows.

However, the system exhibited lower performance in recognizing anger expressions, facing maximum classification errors between anger and sadness due to their subtle and similar changes.

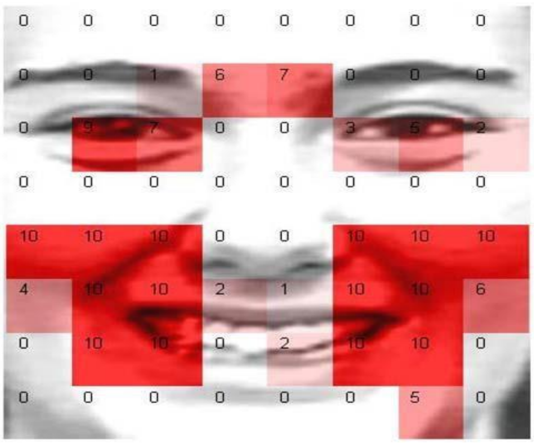



**Fig. 9.** Distribution of selected common patches on faces.The darker the red color is, the more times (shown as numbers) the patch has been selected as common patchesin ten-fold experiments.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 01, January 2024, pp : 196-206

e-ISSN : 2583-1062

Impact Factor : 5.725

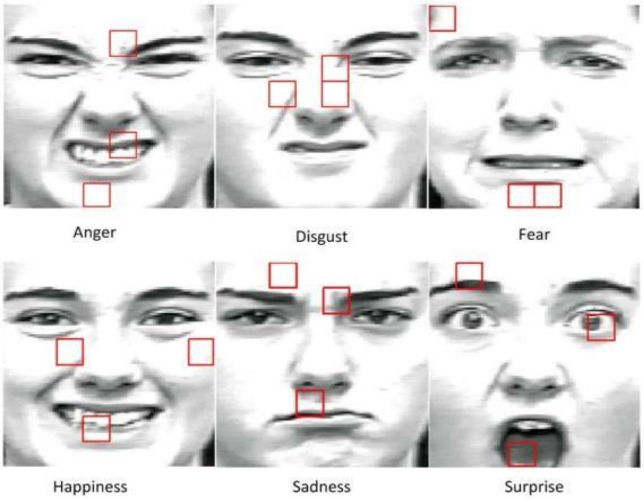



**Fig. 10.** Top three specific patches for six expressions after eliminating the shared patches on the Cohn–Kanade database.

**Optimum Salient Patches -** The number of patches utilized significantly impacted speed and accuracy. Experimentation illustrated that employing features from all 19 patches resulted in an accuracy of 93.87%. Even utilizing appearance features from a single salient patch efficiently discriminated between expression pairs with a recognition rate of 91.19%. Consequently, selective utilization of some salient facial patches significantly enhanced computational complexity and feature robustness, particularly under partial occlusions. Utilizing the top four salient patches achieved close to 95% accuracy.

**Performance Comparison-** The proposed method's performance was compared with similar protocols in the CK+ dataset. While Uddin et al. reported the highest recognition performance for disgust, fear, and happiness, the proposed system achieved an average recognition rate of 94.09%. The use of specific facial patches without temporal features led to high recognition rates, emphasizing the discriminative power of the chosen patches.

**Specific Patches and Multiscale Strategy-** Further analysis revealed the efficacy of integrating specific patches for improved performance. Top three learned specific patches for each expression illustrated their relevance to particular facial features related to the expressions. Integrating these specific patches with common ones via the two-stage MTSL approach amplified the system's recognition rates. The experiment extended to multiscale patches, showcasing the effectiveness of patches around the mouth and eyes across different scales for expression recognition. The incorporation of multiscale patches exhibited better performance than utilizing common patches from a single scale, underscoring the informativenature of multiscale strategies for expression recognition.

**Results on Different Databases-** The evaluation extended to databases like MMI and GEMEP-FERA2011, where the system showcased varying performances due to database-specific challenges. Despite the inherent complexities, the system's performance surpassed comparable methods, underscoring the robustness of the proposed method across different datasets.

proposed method was demonstrated across varied facial

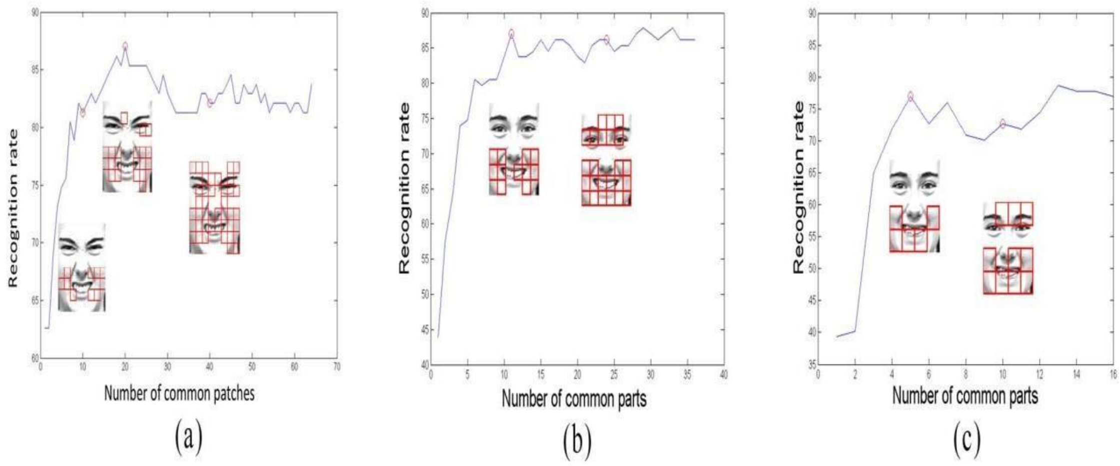


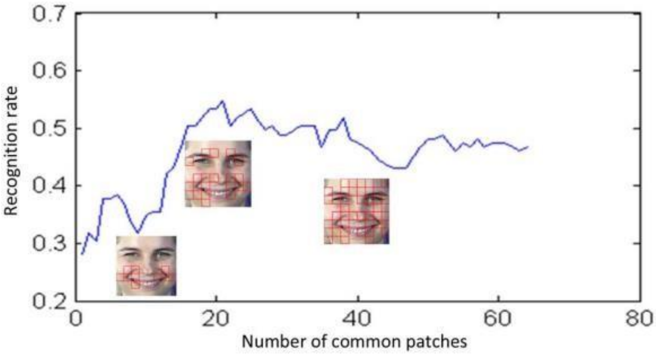

**Fig. 8.** Expression recognition rate with a different number of common patches. (a) Recognition result with selected common patches for the scale S8. The

patch number for the three faces images marked with selected common patches are 10, 20, and 40, respectively.

(b) Results for S6. Patch number are 11 and

24, respectively. (c) Results for S4. Patch number are 5 and 10, respectively. All results show the most effective patches are around the mouth and the eyes,

and using only one third of all the patches can achieve satisfied performance.




Recognition rate with different common patch number. Result ofone fold experiment is shown.

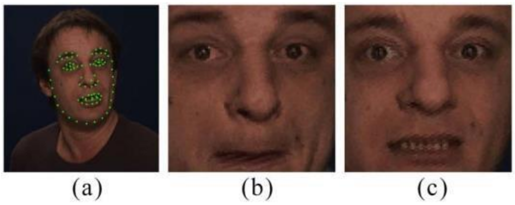



**Fig. (a)** Landmark detection result of [71]. (b) and (c) Aligned andcropped face image examples from the GEMEP-FERA2011 Database.

**Conclusion from Experimental Results-** The experimental outcomes validate the efficacy of the proposed method in recognizing facial expressions across various datasets. Leveraging selective salient patches, integrating specific patches, and adopting multiscale strategies significantly improved recognition rates across different expressions and databases. The robustness of the expressions, offering promise for real-world applications in affective computing and computer vision.

## 6. CONCLUSION

Facial expression analysis and image forgery detection represent pivotal domains in computer vision, each with its unique challenges and advancements. The culmination of these research endeavors culminates in a more comprehensive understanding of these fields, unveiling innovative methodologies and promising avenues for future exploration. The proposed facial expression analysis method offers a paradigm shift in understanding human emotions through facial cues. Departing from traditional approaches, this method delves into the commonalities shared among expressions, unraveling both common and specific facial patches. Leveraging a two-stage sparse learning model and a multiscale face division strategy, this approach mines crucial information from facial muscles and Action Units (AUs). The empirical validation underscores the discriminative power of common patches in expression recognition, augmented further by the integration of specific patches. Additionally, the utilization of multiscale patch division strategies illuminates the diverse coverage areas of facial patches, enhancing recognition performances across different expressions. Conversely, image forgery detection represents a critical facet in safeguarding digital content. The focus on copy– move forgery detection elucidates the challenge posed by this primitive yet intricate form of image manipulation. Detailed exploration of state-of-the-art detection algorithms, encompassing their implementation, performance evaluation, and comparison, marks a significant stride in this domain. The introduction of standardized parameters for performance evaluation and comparison charts a structured path for users to select suitable forgery detection algorithms, considering varied requirements and expected forgery types. Integrating these two diverse domains brings forth a synthesis that underscores the multifaceted nature of computer vision applications. The amalgamation of methodologies from facial expression analysis and image forgery detection offers a holistic perspective on leveraging computer vision in understanding human emotions and protecting digital integrity. The union of these divergent paths underscores the essence of innovation in computer vision research. Drawing parallels between these seemingly disparate domains, parallels emerge. The quest for commonalities in facial expressions mirrors the need to identify consistent features amidst various forgery types. Both domains benefit from detailed exploration, iterative improvements, and a thirst for deeper understanding. As the research landscape unfolds, the paths forwardbecome clearer. Future strides in facial expression analysis may involve further incorporation of specific patches and multiscale strategies, embracing novel computational methodologies and refining recognition systems' robustness. Similarly, the trajectory for image forgery detection points toward advanced training schemes resilient against diverse forgery scenarios and the creation of comprehensive datasets for rigorous evaluation and future research. The synergistic alliance between these domains epitomizes the crossroads where technological advancement meets human-centric understanding. Bridging the gap between interpreting facial cues and safeguarding digital content elucidates the evolving landscape of computer vision's impact on human lives. As the domains continue to evolve, propelled by innovation and interdisciplinary collaborations, they pave the way for a future where technology augments human understanding and secures digital integrity. In conclusion, the synthesis of these divergent yet complementary domains reinforces the multifaceted nature of computer vision. The journey through facial expression analysis and image forgery detection elucidates not only the technical intricacies but also the broader implications for human-computer interaction, emotion understanding, and content integrity. The convergence of these realms represents a promising trajectory for computer vision, propelling it toward new frontiers that amalgamate technological prowess with human-centric applications.

## 7. ROBUSTNESS EVALUATIONS

**1. Dataset Selection and Preprocessing**

Facial Expression Analysis: Utilize benchmark datasets like CK+, MMI, or FERA to ensure standardization and comparability across methodologies. Preprocess data for consistency in resolution, alignment, and illumination. Image Forgery Detection: Employ a diverse set of image datasets encompassing varied types of forgeries (e.g., copy–move, splicing) and realistic scenarios. Uniformly preprocess images to mitigate biases.

**2. Performance Metrics**

Facial Expression Analysis: Employ standard metrics likeprecision, recall, F1-score, and accuracy for expression classification. Utilize confusion matrices to understand classification nuances across different expressions. Image Forgery Detection: Metrics like True Positive Rate (TPR), False Positive Rate (FPR), Receiver Operating Characteristic (ROC) curves, and Precision-Recall curvesoffer insights into forgery detection performances.

**3. Baseline Models and Benchmarking**

Facial Expression Analysis: Compare proposed methodologies against established baseline models, such as traditional machine learning classifiers (SVM, k-NN), to showcase improvements. Benchmark against state-of-the-art methods. Image Forgery Detection: Contrast proposed detection algorithms against baseline approaches like correlation- based detection or block-based methods. Compare performance against contemporary forgery detection techniques.

**4. Cross-Validation and Generalization**

Facial Expression Analysis: Perform k-fold cross- validation to assess model generalizability and robustnessacross different subsets of data. Examine performance variations and ensure consistency. Image Forgery Detection: Test methodologies across diverse datasets to evaluate generalization capacity. Employ cross-dataset validation to gauge performance under varied forgery scenarios.

**5. Computational Efficiency and Scalability**

Measure computational complexities, execution times, and resource requirements (e.g., memory, processing power) of the proposed methodologies against baseline models. Evaluate scalability concerning dataset sizes.

**6. Sensitivity Analysis and Robustness Testing**

Facial Expression Analysis: Analyze model sensitivity to noisy or occluded data by introducing perturbations or partial face occlusions. Assess the robustness of patch- based methods against variations in illumination and pose.

Image Forgery Detection: Conduct sensitivity tests by introducing varied forgery intensities and scales. Evaluate the robustness of forgery detection against compression artifacts or post-processing operations.

**7. Statistical Significance and Conclusion**

Apply statistical tests (e.g., t-tests) to determine the significance of observed differences in performance metrics between methodologies.

Conclude with a comprehensive comparative analysis, emphasizing the strengths and limitations of reviewed methodologies in facial expression analysis and image forgery detection. Although the proposed scheme is mainly designed to counter the lossy operations conducted by OSNs, and also like to evaluate its robustness under some more commonly used degradation scenarios, such as noise addition,cropping, resizing, blurring, and standalone JPEG compression. Such evaluation is very critical in real-world cases because these types of post- processing operations are often adopted to erase or conceal the forged artifacts. To this end, applying these post-processing operations to the original test set Columbia and report the quantitative comparisons. For the convenience of demonstration, utilizing a unified parameter $p$ for controlling the magnitudes of different operations. The origin of the horizontal axis ($p$ = 0) corresponds to the case without any postprocessing. It can be observed, the competitors [12,27] cannot perform consistently with the increase of the perturbation intensity, while this method can generalize well to defeat these post processing operations.

## 8. REFERENCES

[1] (2013, Oct. 8). Facial expression [Online]. Available:http://en.wikipedia.org/wiki/Facial_expression

[2] T. G. Dietterich, L. Pratt, and S. Thrun, "Special issue on inductive transfer," Mach. Learn., vol. 28, no. 1, Jul. 1997.

[3] A. Argyriou and T. Evgeniou, "Multi-task feature learning," in Advances in Neural Information Processing Systems, vol. 19.Cambridge, MA, USA: MIT Press, 2007.

[4] T. Bänziger, M. Mortillaro, and K. R. Scherer, "Introducing the Geneva multimodal expression corpus for experimental research on emotion perception," Emotion, vol. 12, pp. 1161–1179, 2012.

[5] T. Bänziger and K. R. Scherer, "Introducing the Geneva multimodal emotion portrayal (GEMEP) corpus," in Blueprint for Affective Computing: A Sourcebook, K. R. Scherer, T. Banziger, and E. B. Roesch, Eds. Oxford, U.K.: Oxford Univ. Press, 2010, pp. 271–294.

[6] M. Bartlett, J. Hager, P. Ekman, and T. Sejnowski, "Measuring facial expressions by, compute image analysis," Psychophysiology, vol. 36, no. 2, pp. 253–263, Mar. 1999.

[7] M. S. Bartlett et al., "Recognizing facial expression: Machine learning and application to spontaneous behavior," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2. Jun. 2005, pp. 568–573.

[8] M. S. Bartlett et al., "Fully automatic facial action recognition in spontaneous behavior," in Proc. 7th Int. Conf. Autom. Face Gesture Recognit., Southampton, U.K., 2006, pp. 223–230.

[9] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Trans. Inf. Theory, vol. 52, no. 2, pp. 489–509, Feb.2006.

[10] R. Caruana, "Multi-task learning," Mach. Learn., vol.28, no. 1, pp. 41–75, 1997.

[11] Y. Chang, C. Hu, R. Feris, and M. Turk, "Manifold based analysis of facial expression," Image Vis. Comput., vol. 24, no. 6, pp.605–614, Jun. 2006.

[12] J. Chen, J. Liu, and J. Ye, "Learning incoherent sparse and low-rank patterns from multiple tasks," ACM Trans. Knowl. Dis.Data, vol. 5, no. 4, Feb. 2012, Art ID. 22.

[13] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," SIAM Rev., vol. 43, no. 1, pp. 129–159, 2001.

[14] S. W. Chew et al., "Person-independent facial expression detection using constrained local models," in Proc. IEEE Int. Conf. Autom. Face Gesture Recognit. Workshops, Santa Barbara, CA, USA, 2011, pp.915–920.

[15] J. F. Cohn, "Foundations of human computing: Facial expression and emotion," in Proc. Int. Conf. Multimodal Interfaces, 2006, pp. 223–238.

[16] J. F. Cohn, L. Reed, Z. Ambadar, J. Xiao, and T. Moriyama, "Automatic analysis and recognition of brow actions and head motion in spontaneous facial behavior," in Proc. IEEE Int. Conf. Syst., Man, Cybern., 2004, pp. 610–616.

[17] J. F. Cohn and K. L. Schmidt, "The timing of facial motion in posed and spontaneous smiles," Int. J. Wavelets Multiresolut. Inf.Process., vol. 2, no. 2, pp. 1–12, 2004.

[18] G. Duchenne, Mecanisme de la Physionomie Humaine. Paris, France: Renouard, 1862.

[19] P. Ekman, W. V. Friesen, and J. C. Hager, Facial Action Coding System: A Technique for the Measurement of Facial Movement. Palo Alto, CA, USA: Consulting Psychologists Press, 2002.

[20] I. Essa and A. Pentland, "A vision system for observing and extracting facial action parameters," in Proc. IEEE Comput. Soc.Conf. Comput. Vis. Pattern Recognit. (CVPR), Seattle, WA, USA, 1994, pp. 76–83.

[21] B. Fasel and J. Luettin, "Automatic facial expression analysis: A survey," Pattern Recognit., vol. 36, no. 1, pp. 259–275, Jan. 2003.

[22] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," IEEE J. Sel. Topics Signal Process., vol. 1, no. 4, pp. 586–597, Dec. 2007.

[23] G. Obozinski, "Multi-task feature selection," Dept. Statistics, Univ. California, Berkeley, Berkeley, CA, USA, Tech. Rep.,2006.

[24] G. Guo and C. R. Dyer, "Learning from examples in the small sample case—Face expression recognition," IEEE Trans. Syst.,Man, Cybern. B, Cybern., vol. 35, no. 3, pp. 477–488, Jun. 2005.

[25] J. Huang, X. Huang, and D. Metaxas, "Learning with dynamic group sparsity," in Proc. IEEE 12th Int. Conf. Comput. Vis.,Kyoto, Japan, 2009, pp. 64–71.

[26] C. E. Izard, The Face of Emotion, vol. 1. New York, NY, USA: Appleton- Century-Crofts, 1971.

[27] B. Jiang, M. F. Valstar, and M. Pantic, "Action unit detection using sparse appearance descriptors in space-time," in Proc. IEEE Int. Conf. Autom. Face Gesture Recognit. Workshops, Santa Barbara, CA, USA, 2011, pp. 314–321.

[28] T. Kanade, J. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," in Proc. 4th IEEE Int. Conf. Autom.Face Gesture Recognit., Grenoble, France, 2000, pp. 46–53.

[29] S. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interiorpoint method for large-scale l1-regularized least squares," IEEE J. Sel. Topics Signal Process., vol. 1, no. 4, pp. 606–617, Dec.2007.

[30] G. Littlewort, M. S. Bartlett, J. S. I. Fasel, and J. Movellan, "Dynamics of facial expression extracted automatically from video,"in Proc. Int. Conf. Comput. Vis. Pattern Recognit., 2004, p. 80.

[31] M. Lyons, J. Budynek, and S. Akamatsu, "Automatic classification of single facial images," IEEE Trans. Pattern Anal. Mach.Intell., vol. 21, no. 12, pp. 1357–1362, Dec. 1999.

[32] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," IEEE Trans. Signal Process., vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[33] M. F. Valstar, B. Jiang, M. Mehu, M. Pantic, and K. Scherer, "The first facial expression recognition and analysis challenge," inProc. IEEE Int. Conf. Autom. Face Gesture Recognit., Santa Barbara, CA,USA, 2011, pp. 921–926.

[34] M. F. Valstar, B. Jiang, M. Mehu, M. Pantic, and K. Scherer, "Metaanalysis of the first facial expression recognition challenge," IEEETrans. Syst., Man, Cybern. B, Cybern., vol. 2, no. 4, pp. 966–979, Aug. 2012.

[35] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with localbinary patterns," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp.971–987, Jul. 2002.

[36] M. Pantic and T. Patras, "Dynamics of facial expression: Recognition of facial actions and their temporal segments form face profile image seqences," IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 36, no. 2, pp. 433–449, Apr. 2006.

[37] M. Pantic and L. J. M. Rothkrantz, "Case-basedreasoning for userprofiled recognition of emotions from face images," in Proc. IEEEInt. Conf. Multimedia Expo, Taipei, Taiwan, 2004, pp. 391–394.

[38] A. Ryan et al., "Automated facial expression recognition system," in Proc. 43rd Annu. Int. Carnahan Conf. SecurityTechnol., Zurich, Switzerland, 2009, pp. 172–177.

[39] G. Sandbach, S. Zafeiriou, and M. Pantic, "Binary pattern analysis for 3D facial action unit detection," in Proc. Brit. Mach. Vis.Conf., Surrey, U.K., 2012, pp. 119.1–119.12.

[40] G. Sandbach, S. Zafeiriou, M. Pantic, and D. Rueckert, "Recognition of 3D facial expression dynamics," Image Vis. Comput., vol. 30, no. 10, pp. 762–773, Oct. 2012. Authorized licensed use limited to: VTU Consortium. Downloaded on December 11,2023 at 10:59:39 UTC from IEEE Xplore. Restrictions apply. 1510 IEEE TRANSACTIONS ON CYBERNETICS, VOL. 45, NO. 8, AUGUST 2015

[41] N. Sebe et al., "Authentic facial expression analysis," in Proc. 6th IEEE Int. Conf. Autom. Face Gesture Recognit., 2004, pp. 517–522.

[42] T. Senechal et al., "Facial action recognition combining heterogeneous features via multikernel learning," IEEE Trans. Syst.,Man, Cybern. B, Cybern., vol. 42, no. 4, pp. 993–1005, Aug. 2012.

[43] C. Shan, "Smile detection by boosting pixel differences," IEEE Trans. Image Process., vol. 21, no. 1, pp. 431–436, Jan. 2012.

[44] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study,"Image Vis. Comput., vol. 27, no. 6, pp. 803–816, May 2009.

[45] Y. Tian, "Evaluation of face resolution for expression analysis," in Proc. Int. Conf. Comput. Vis. Pattern Recognit., Jun. 2004, p.82.

[46] Y. Tian, T. Kanade, and J. F. Cohn, “Recognizingupper face action units for facial expression analysis,” in Proc. Int. Conf.Comput. Vis. Pattern Recognit., 2000, pp. 294–301.

[47] Y. Tian, T. Kanade, and J. F. Cohn, “Recognizingaction unites for facial expression analysis,” IEEE Trans. Pattern Anal. Mach.Intell., vol. 23, no. 2, pp. 97–115, Feb. 2001.

[48] Y. Tong, J. Chen, and Q. Ji, “A unified probabilisticframework for spontaneous facial action modeling and understanding,”IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 2, pp. 258–273,Feb. 2010.

[49] Y. Tong, W. Liao, and Q. Ji, “Facial action unitrecognition by exploiting their dynamic and semantic relationships,” IEEE Trans.Pattern Anal. Mach. Intell., vol. 29, no. 10, pp. 1683–1699, Oct. 2007.

[50] J. A. Tropp and C. Gilbert, “Signal recovery fromrandom measurements via orthogonal matching pursuit,” IEEE Trans. Inf.Theory, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[51] M. Valstar and M. Pantic, “Fully automatic facialaction unit detection and temporal analysis,” in Proc. Int. Conf. Comput. Vis.Pattern Recognit., 2006, p. 149.

[52] M. F. Valstar and M. Pantic, “Induced disgust, happiness and surprise: An addition to the MMI facial expression database,” inProc. Int. Conf. Lang. Resour. Eval., 2010, pp. 65–70.

[53] M. F. Valstar, M. Pantic, Z. Ambadar, and J. F. Cohn,“Foundations of human computing: Facial expression and emotion,” inProc. Int. Conf. Multimodal Interfaces, 2006, pp. 162–170.

[54] A. Vinciarelli, M. Pantic, and H. Bourlard, “Socialsignal processing: Survey of an emerging domain,” Image Vis. Comput., vol.31, no. 1, pp. 1743–1759, 2009.

[55] P. Viola and M. Jones Robust real-time objectdetection Int. J. Comput Vis.vol. 57, no. 2, pp. 137–154, 2001.

[56] E. Vural et al., “Automated drowsiness detection forimproved driver safety comprehensive databases for facial expressionanalysis,” in Proc. Int. Conf. Autom. Technol., vol. 1. Istanbul, Turkey, 2008,pp. 96–105.

[57] X. Wang, C. Zhang, and Z. Zhang, “Boosted multi-task learning for face verification with applications to web image and videosearch in Proc. Int. Conf. Comput. Vis. Pattern Recognit. Miami,FL USA, 2009, pp. 142–149.

[58] J. Whitehill and C. W. Omlin, “Haar features forFACS AU recognition,” in Proc. Int. Conf. Autom. Face Gesture Recognit.,Southampton, U.K., 2006, pp. 97–101.

[59] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma,“Robust face recognition via sparse representation,” IEEE Trans.Pattern Anal. Mach. Intell., vol. 31, no. 2, pp. 210–227, Feb. 2009.

[60] C. Xi, P. Weike, J. T. Kwok, and J. G. Carbonell,“Accelerated gradient method for multi-task sparse learning problem,” in Proc.Int. Conf. Data Min., 2009, pp. 746–751.

[61] P. Yang, Q. Liu, and D. N. Metaxas, “Exploring facialexpressions with compositional features,” in Proc. Int. Conf. Comput.Vis. Pattern Recognit., San Francisco, CA, USA, 2010, pp. 2638–2644.

[62] L. Yin, X. Wei, Y. Sun, J. Wang, and M. J. Rosato, “A 3D facial expression database for facial behavior research,” in Proc. Int. Conf.Autom. Face Gesture Recognit., Southampton, U.K., 2006, pp.211–216.

[63] X. Yuan and S. Yan, “Visual classification with multi- task joint sparse representation,” in Proc. Int. Conf. Comput. Vis. PatternRecognit., San Francisco, CA, USA, 2010, pp. 3493–3500.

[64] S. Zafeiriou and I. Pitas, “Discriminant graph structures for facial expression recognition,” IEEE Trans. Multimedia, vol. 10,no. 8, pp. 1528–1540, Dec. 2008.

[65] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang,“A survey of affect recognition methods: Audio, visual and spontaneousexpressions,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 1, pp. 39–58, Jan. 2009.

[66] S. Zhang et al., “Automatic image annotation usinggroup sparsity,” in Proc. Int. Conf. Comput. Vis. Pattern Recognit., SanFrancisco, CA, USA, 2010, pp. 3312–3319.

[67] S. Zhang, J. Huang, H. Li, and D. Metaxas,“Automatic image annotation and retrieval using group sparsity,” IEEE Trans. Syst.,Man, Cybern. B, Cybern., vol. 42, no. 3, pp. 838–849, Jun. 2012.

[68] S. Zhang, Y. Zhan, M. Dewan, J. Huang, and D. Metaxas,“Towards robust and effective shape modeling: Sparse shapecompositon,” Med. Image Anal., vol. 16, no. 1, pp. 265–277, Jan. 2012.

[69] Y. Zhang and Q. Ji, “Active and dynamic informationfusion for facial expression understanding from image sequences,” IEEETrans. Pattern Anal. Mach. Intell., vol. 27, no. 5, pp. 699–714, May2005.

[70] L. Zhong et al., “Learning active facial patches forexpression analysis,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit.(CVPR), Providence, RI, USA, 2012, pp. 2562–2569.

[71] E. Zhou, H. Fan, Z. Cao, Y. Jiang, and Q. Yin,“Extensive facial landmark localization with coarse-to-fine convolutional networkcascade,” in Proc. IEEE Int. Conf. Comput. Vis.