Facial Action Units Detection using Transfer Learning

line 1: 1st Given Name Surname   
line 2: *dept. name of organization   
(of Affiliation)*  
line 3: *name of organization   
(of Affiliation)*line 4: City, Country  
line 5: email address or ORCID

Sukrit Jaidee  
*Department of Electrical Engineering,  
Faculty of Engineering,*  
*Chulalongkorn University*Bangkok, Thailand  
6170306321@alumni.chula.ac.thline 1: 2nd Given Name Surname  
line 2: *dept. name of organization   
(of Affiliation)*  
line 3: *name of organization   
(of Affiliation)*line 4: City, Country  
line 5: email address or ORCID

line 1: 3rd Given Name Surname  
line 2: *dept. name of organization   
(of Affiliation)*  
line 3: *name of organization   
(of Affiliation)*line 4: City, Country  
line 5: email address or ORCID

*Abstract*— This paper presents a facial unit change detection model using an artificial intelligence model. Key features, including Histogram of Oriented Gradient (HOG), facial landmarks/keypoints, and centre keypoint angles, were extracted as inputs to the model. Moreover, the optimal parameters of the model are determined to optimise the model. This model is trained on the Ck+ database with 99% accuracy, and compared to the base model, it is 8% more accurate.

Keywords— Facial Action Units, Transfer learning, facial keypoint, face landmark, Emotion

# Introduction

It is difficult to estimate the emotional expression of people undergoing psychiatric treatment in the early stages of research. Researchers have been trying to understand how the brain perceives emotions from facial expressions. It is assumed that these are caused by learning to recognise facial movements. In 1978, American psychologists Ekman and Friesen classified facial movements into 46 Action Units (AUs) [1], calling it the Facial Action Coding System (FACS). This system uses a method to number each muscle movement that there are 46 facial parts. This system has since been accepted as the standard in the study of facial expressions [2].

FACS at the beginning of studies and research is limited as assessors must be trained until they are certified. The results are therefore reliable. In 2011, Hamm J developed an algorithm to assess FACS [3]. After that, researchers have studied and researched FACS more widely in behavioural medicine. In the early days after that, FACS was used to categorize emotional expressions.

For example, a happy mood is explained by the FACS that the cheek muscles are lifted. The muscles around the eyes contract, creating a crease in the corner of the eye. The corners of the mouth arched up into a symmetrical smile. Considering the detected units, Cheek Raiser (AU6) and Lip Corner Puller (AU12) in TABLE I. , the faces being assessed were happy, as shown in Figure 2.

2. Sadness Inner Brow Raiser(1), Brow Lowerer(4), Lip Corner Depressor(15) หัวคิ้วของเราจะย่นเข้าหากัน ทำให้เกิดริ้วย่นรูปตัวยู่คว่ำ และเกิดร่องแนวตั้งที่หว่างคิ้ว ทั้งหมดนี้เรียกว่า "กล้ามเนื้อเศร้าของดาร์วิน" มุมปากด้านนอกจะตกลง 3. Anger Brow Lowerer(4), Upper Lid Raiser(5), Lid Tightener(7), Lip Tightener(23) หัวคิ้วจะย่นเข้าหากัน ทำให้เกิดริ้วรอยในแนวตั้งที่เห็นได้ชัดเจน ซึ่งเรียกกันว่า "คิ้วพันกัน" นอกจากนี้ยังอาจทำให้เปลือกตาด้านบนเบิกขึ้น ทำให้เห็นตาขาว และเกิดการแสดงออกที่เครียดเขม็ง เมื่อเปลือกตาล่างเกร็งเข้าหากัน ริมฝีปากเม้มแน่นและมองไม่เห็นริมฝีปากบนในส่วนที่เป็นสีชมพู 4. Fear Inner Brow Raiser (1), Outer Brow Raiser(2), Brow Lowerer(4), Upper Lid Raiser(5), Lid Tightener(7), Lip Stretcher(20), Jaw Drop(26) คิ้วของเราจะแบนเกือบราบ เกิดรอยย่นบนเส้นคิ้ว เช่นเดียวกับความโกรธ ตาจะเบิกกว้าง เปลือกตาบนเปิด ทำให้มองเห็นตาขาว มุมปากฉีกออกไปด้านข้างเป็นเส้นตรงแน่น 5. Surprise Inner Brow Raiser(1), Outer Brow Raiser(2), Upper Lid Raiser(5), Jaw Drop(26) เมื่อเราประหลาดใจ เปลือกตาบนจะถูกยกขึ้น ทำให้เห็นตาขาว และอ้าปากหรือขากรรไกร 6. Disgust Nose Wrinkler(9), Lip Corner Depressor(15), Lower Lip Depressor(16) เราจะย่นจมูกเหมือนเวลาได้กลิ่นเหม็น ทำให้เกิดรอยย่นแนวนอนที่บริเวณจมูกใกล้หว่างคิ้ว มีการแบะริมฝีปากบนร่วมด้วย 7. Contempt Lip Corner Puller(12), Dimpler(14) มุมปากซ้ายจะฉีกออกด้านข้าง เกิดเป็นรอยบุ๋ม และโดยมากจะกลอกตาร่วมด้วย

From the above example¸ Most research today uses action units to interpret emotions which is a description of the emotion in the term of the unit action as shown in TABLE I.

1. Emotion description in terms of action units.

| **Emotion** | Action Unit Criteria |
| --- | --- |
| Angry | 4, 5, 7, 23, 24 (optional) |
| stress | 1,17,25,26 |
| Disgust | 9 or 10, 15, 16 |
| Fear | 1+2+4+5+7+20+26 |
| Happy | 6+7+12+25 |
| Sadness | 1+4+15 |
| Surprise | 1 + (2 or 5) + 26 |
| relax | 12,25,26,43(optional) |
| Contempt | 12, 14 |

# Related Work

After researching several articles, article [5] presents a model for detecting action units by training the model on the ck+ database. This paper uses a Local Binary Pattern (LBP) to extract key features and textures of an image and has reduced its dimensions using PCA, LDA, Adaboost, and Gentleboost to train SVM classifiers with RBF/Linear kernel and employ hyperparameter tuning techniques.

# FACIAL ACTION UNITS DeTEction

## Action unit detection process overview

Action unit detection consists of 3 steps: 1. Face detection 2. Extract facial keypoints or landmarks from the face. 3. Estimate the change in facial keypoints or landmarks as shown in Figure 1.

|  |
| --- |
|  |

1. Action unit detection process overview

The templat

## Face detection

การตรวจจับภาพใบหน้าปัจจุบันมีหลากหลายวิธีการซึ่งมีข้อดีข้อเสียแตกต่างกันไปทั้งในแง่ของความแม่นยำและความเร็วในการประมวลผล ยกตัวอย่างเช่น วิธีการแรก เป็นการแยกใบหน้าออกจาก background มี feature input เป็นการเปลี่ยนแปลงไล่ระดับของสีในภาพ (Histogram of Oriented Gradient: HOG) และโมเดลที่ใช้ในการ classify ส่วนใบหน้า และ background คือ Support Vector Machine (SVM) ซึ่งวิธีการนี้มีข้อเสียตรงที่การถ่ายภาพในที่มีแสงน้อยอาจมีผลกระทบต่อความแม่นยำได้ วิธีการต่อมาคือ … เป็นต้น

ปัจจุบันมี libraries สำหรับการ face detection ในหลายๆโมเดลซึ่งในแต่ละ library จะมีความแตกต่างกันในสองปัจจัยหลักก็คือ acc และ time ดังแสดงในตารางที่ x ด้านล่าง เครื่องหมายถูกหมายถึงมีจุดเด่นในด้านนั้น

1. Face detection

| No | Face detection | | |
| --- | --- | --- | --- |
| Libraries | Time | acc |
| 1 | Dlib | ✓ |  |
| 2. | MTCNN |  |  |
| 3. | FaceBoxes |  |  |
| 4. | RetinaFace | ✓ | ✓ |
| 5. | Haar cascade |  |  |

## Facial landmark

การประเมินการขยับของกล้ามเนื้อใบหน้าหรือการอธิบายการเคลื่อนที่ส่วนต่างๆของใบหน้าต้องอาศัยการกำหนดจุดอ้างอิง (landmark) และการกำหนด action unit ต้องอาศัยจุดอ้างอิงนั้น ดังนั้น facial landmark จึงเป็นหนึ่งใน feature ที่สำคัญสำหรับการตรวจจับหน่วยกระทำ ปัจจุบันมี libraries สำหรับการ detect face landmark ในหลายๆโมเดลซึ่งในแต่ละ library จะมีความแตกต่างกันในสามปัจจัยหลักก็คือ acc , size และ time ดังแสดงในตารางที่ x ด้านล่าง เครื่องหมายถูกหมายถึงมีจุดเด่นในด้านนั้น

1. Face Landmark

| No | Face Landmark | | | | |
| --- | --- | --- | --- | --- | --- |
| Libraries | Time | acc | size |
| 1 | Dlib | ✓ |  | ✓ |
| 2 | Openpose |  | ✓ |  |
| 3 | Mediapipe | ✓ | ✓ | ✓ |
| 4 | Pre-Trained model  (Xception, ResNet,  VGG) |  | ✓ |  |
| 5 | stasm | ✓ |  | ✓ |
| 6 | PFLD |  |  |  |
| 7 | MobileFaceNet |  |  | ✓ |
| 8 | MobileNet |  |  |  |

สำหรับตัว Open pose เราทดสอบการทดสอบทั้งที่เป็น State of the art, Caffe, TensorFlow และ Chainer model

## Action Unit Detection (classification)

การตรวจจับ action unit สามารถตรวจจับได้จากการเปลี่ยนของ face landmarks ต่างๆบนใบหน้า ปัจจุบันมี libraries สำหรับการตรวจจับ action unit อยู่ไม่กี่โมเดล ดังแสดงในตารางที่ x ซึ่งปัจจุบันยังมีความแม่นยำที่ไม่สูงมาก จากการศึกษาพบว่า หลายๆงานวิจัยกำลังปรับปรุงความแม่นยำของการตรวจจับ action unit

วิธีการที่นำเสนอได้สร้าง feature ขึ้นมาใหม่สอง feature ได้แก่ ระยะห่างระหว่างแต่ละ face landmark กับ center point ของทุก landmark และ มุมของแต่ละ face landmark กับ center ของทุก face landmark ซึ่ง feature ทั้งหมดที่นำเสนอจะอยู่ในรูป <x, y, distance from center, angle relative to center> ซึ่ง feature เหล่านี้จะช่วยลดความผิดผลาดที่เกิดจากการเอนเอียงของใบหน้า

1. Action unit detection

| No | Action unit detection | | | | |
| --- | --- | --- | --- | --- | --- |
| Libraries | Time | acc | size |
| 1 | random forest |  |  |  |
| 2 | svm |  |  |  |
| 3 | logistic |  |  |  |
| 4 | JAANET |  |  |  |
| 5 | DRML |  |  |  |
| 6 | Perposed model |  |  |  |

## Regression model

Regression เป็นการบอกว่า frame นั้นมีความเปลี่ยนแปลงไปจาก frame ก่อนในระดับมากน้อยเพียงใด

# อธิบายข้อมูล

Dataset ที่ใช้ในงานวิจัยนี้คือ CK+ dataset เนื่องมี labelled AUs และ labelled basic emotions. Dataset นี้ประกอบไปด้วย 593 image sequences ของ 123 subjects. อนุกรมของรูปภาพทั้งหมดเริ่มที่การแสดงออกทางอารมณ์แบบปกติและสิ้นสุดด้วยการแสดงออกทางอารมณ์แบบ peak โดยการ label AU นั้นมีการ label 2 รูปแบบ ได้แก่ 1. Label ที่บอกว่า AU นั้นเกิดขึ้นหรือไม่ (classification) 2. Label ที่บอกความรุ่นแรงของ AU ที่เกิดขึ้น (regression)

มี 327 อนุกรมรูปภาพที่ถูก label สำหรับ basic emotions.

## Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even.

## Units

* Use either SI

## Equations

The equations are an exception to the prescribed specifications of

*a**b* 

Note that the equation is centered using a center tab stop. Be sure that

## Some Common Mistakes

* The word “data” is plural, not singular.
* The subscript for the permeability of vacuum **0, and other

An excellent style manual for science writers is [7].

# การทดลอง

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

## Face detection

**The template is designed for, but not limited to, six authors.** A minimum of one author is required for all conference article

### For papers with more than six authors: Add author names horizontally, moving to a third row if needed for more than 8 authors.

### For papers with less than six authors: To change the default, adjust the template as follows.

#### Selection: Highlight all author and affiliation lines.

#### Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

#### Deletion: Delete the author and affiliation lines for the extra authors.

## Facial Landmark detection

ผลการทดลองจะเป็นการเปรียบเทียบ acc และ time แล้วเลือก face landmark model ที่เหมาะสมที่สุด

## Action Units Detection

ผลการทดสอบการตรวจจับ action unit บน CK+ dataset แสดงดังตารางด้านล่าง

#### Positioning Figures and Tables: Place figures and tables at the top

1. Table Type Styles

| Table Head | Table Column Head | | |
| --- | --- | --- | --- |
| Table column subhead | Subhead | Subhead |
| copy | More table copya |  |  |

1. Sample of a Table footnote. (*Table footnote*)
2. Example of a figure caption. (*figure caption*)

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis

# ผลลัพธ์และสรุป

แสดง heatmap layer สุดท้ายของ model

แสดงกราฟ time-series ของ action unit

##### Acknowledgment *(Heading 5)*

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

# References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

1. https://imotions.com/blog/facial-action-coding-system/
2. Facial Action Coding System [Internet]. Paul Ekman Group. [cited 2020 Jun 28]. Available from: <https://www.paulekman.com/facial-action-coding-system/>
3. Hamm J, Kohler CG, Gur RC, Verma R. Automated Facial Action Coding System for dynamic analysis of facial expressions in neuropsychiatric disorders. Journal of Neuroscience Methods 2011;200:237-56.
4. Kunz M, Seuss D, Hassan T, Garbas JU, Siebers M, Schmid U, et al. Problems of video-based pain detection in patients with dementia: a road map to an interdisciplinary solution. BMC Geriatrics [Internet]. 2017 Dec [cited 2018 Mar 22];17(1). Available from: http://bmcgeriatr.biomedcentral.com/articles/10.1186/s12877-017-0427-2
5. Y. Zhang, W. Dong, B. Hu and Q. Ji, "Classifier Learning with Prior Probabilities for Facial Action Unit Recognition," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 5108-5116, doi: 10.1109/CVPR.2018.00536.
6. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
7. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
8. K. Elissa, “Title of paper if known,” unpublished.
9. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
10. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
11. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.