# Using Machine Learning Algorithms to Detect Smile in Pictures Based on Geometric Features

**Prepared by:**

**Doa'a Ahmad Jebreel Mohammad**

**Supervised by:**

**Dr. Mohammad Othman Nassar**

## Abstract:

Usage of Digital images applications have become one of the most used applications these days. It is increased rapidly and developers are working continuously to improve them by adding new features to make them easier to use and more user friendly. One of the most required features in those applications is smile detection.

Searching literature and researches about digital smile detection revealed a lot of methods. However, none of these applies angles based approach to detect smile in images using machine learning software.

This research is devoted to detect smile in pictures passing through stepwise stages. first stage use Flandmark algorithm used to detect human

faces in pictures, followed by detecting human lips by using the same algorithm.

In the second stage, lips landmarks are determined and used to draw lines and generate five angles for the lips. Those angles are used create new parameter dataset that's will be the first of its kind. The third stage has three separated levels; according to the number of angles to be used, level one uses two angles (horizontal angles) with/without the distance between points. In the second level, three angles were used also with/without the distance between points, and finally we will work with five angles in the third level.

At the third stage, we create the new dataset from the previous angles by changing it to parameters form with (. arff) file type. At the fourth stage, we divide our method into many probabilities to determine the best probability to detect the smile; these probabilities depend on the number of parameter that used in each experiment: Using ang1_ang2 and the two distances, using ang1_ang3 and the two distances, using ang2_ang4 and the two distances, using ang3_ang4_ang5 and the distances, using all angles and the distances, and finally; Using all angles without distances.

all probabilities are tested three times, the first and the second will be tested by cross validation method using many Weka classifiers with modifications and without modifications on classifiers properties, And the third one by using training/testing method with properties modifications.

In this research, nine classifiers were used: Bayes Net, Naïve Bayes, SMO, IBK, Decision Table, BF Tree, J48, Random Tree, and Rep Tree. The number of experiments reached to 594 to include all factors that can affect the results and accuracy of the research.

The results that we get indicates that our proposed method is more efficient than 2DPCA, Adaboost, PCA and Shan et al., and give us 100% accuracy when using IBK and Random tree classifiers with training/testing method with any combination of angles. Also we found that SMO classifier in most cases can detect all non-smiley faces in the dataset so it's the best classifier to be used to detect non-smile faces but cannot detect any smiley face.

# استخدام التعلم الالي لرصد الابتسامة اعتمادا على الخصائص الهندسية لشفاه الفم

**اعداد**

**دعاء احمد جبريل محمد**

**اشراف**

**د. محمد عثمان نصار**

## الملخص

ازداد استخدام الصور الرقمية وتطبيقاتها في الأعوام الأخيرة بشكل ملحوظ حيث أصبحت تدخل في شتى مجالات ومناحي الحياة اليومية وأصبحت جزء لا يتجزأ من اعملنا ومهامنا. حيث أصبحت تستخدم لحفظ مذكراتنا وتفاصيل أيامنا، من هنا فقد أصبحت الحاجة لتطوير تطبيقات الصور الرقمية واستخداماتها امرا مهما.

من اهم تطبيقات الصور الرقمية هي استخدامها لأخذ صور الشخصية للأفراد، بحيث تظهر تعابير الوجه للأفراد والتي يمكن من خلالها استنتاج الحالة المزاجية وشعور الشخص عند التقاط الصورة. ولهذا السبب فانه سيتم البحث في هذا المجال ومحاولة استخدام طرق جديدة في هذا البحث لمعرفة وتحديد بعض تعابير الوجه للإنسان اثناء التقاط الصورة.

تهدف هذه الدراسة الى تقديم الية جديدة يمكن من خلالها تحديد إذا ما كان الشخص في الصورة مبتسما ام غير مبتسم اثناء التقاط الصورة من خلال مجموعة من المراحل وتطبيق هذه المراحل على مجموعة من الصور.

في المرحلة الأولى: يتم استخدام خوارزمية Flandmark لتحديد الوجوه في الصور ومن ثم تحديد الشفاه ايضا باستخدام الخوارزمية نفسها.

في المرحلة الثانية: يتم تحديد نقاط الشفاه ورسم الخطوط بينها لإنتاج خمسة زوايا واستخدامها لإنشاء قاعدة بيانات جديدة من نوعها حيث انها تحتوي على متغيرات وتفاصيل للفم لم يسبق ان تم انتاجها واستخدامها من قبل وستكون متاحة للاستخدام في أبحاث أخرى. يمكن تقسيم هذه المرحلة الى ثلاثة مستويات منفصلة اعتمادا على عدد الزوايا في كل مستوى؛ حيث يحتوي المستوى الأول على زاويتين فقط. بينما يشتمل المستوى الثاني على ثلاثة زوايا، اما المستوى الثالث فيستخدم خمسة زوايا وجميع هذه المستويات لها احتمالين وهما استخدام هذه الزوايا مع مسافتين بين نقاط الفم والاحتمال الثاني هو عدم استخدام هذه المسافات.

بعد تحديد احتمالات الدراسة يتم استخدام وتطبيق بعض الخوارزميات الخاصة بالتصنيف من برمجية ال WEKA على هذه الاحتمالات ومقارنة النتائج لجميع هذه الخوارزميات والاحتمالات. حيث سيتم استخدام الخوارزميات ( Bayes Net, Naïve Bayes, SMO, IBK, Decision Table, BF Tree, J48, Random Tree, Rep Tree). بحيث يتم استخدامها بشكل مباشر دون تعديل خصائصها ومن ثم إعادة استخدامها مع تعديل الخصائص الخاصة بكل خوارزمية لمراقبة ومتابعة التأثير على النتائج لهذه الخوارزميات.

تم اجراء العديد من التجارب التي تم استخدام العديد من الاحتمالات فيها لتصبح شاملة لجميع العوامل التي يمكن ان تؤثر في نتائج الدراسة ودقتها ولتحديد أفضل الطرق لاكتشاف الابتسامة في الصور الرقمية حيث وصل عددها الى 594 تجربة باستثناء التجارب التي استخدمت لتحديد أفضل التعديلات على كل خوارزمية.

تبين في نهاية هذه الدراسة من خلال اجراء بعض المقارنات بين جميع النتائج لجميع الاحتمالات السابقة ان أفضل الطرق المستخدمة للكشف عن الابتسامة في الصور هي طريقة IBKاو الطريقة التي تستخدم Random Tree بغض النظر عن الاحتمال المستخدم من الزوايا لان جميع الاحتمالات باستخدام هذه طريقة تعطي نتيجة اكتشاف بنسبة 100% بشرط استخدامها بأسلوب Training/testing.

وتبين ان المصنف SMOهو أفضل الخوارزميات التي يمكن استخدامها للكشف عن الوجوه غير المبتسمة في اغلب الأحيان حيث تمكن من اكتشاف جميع الصور غير المبتسمة بينما كان الأسوأ في اكتشاف الصور المبتسمة.