

## 台灣情緒語料庫建置與辨識

Bo-Chang Chiou and Chia-Ping Chen

### 摘要

現在已有許多公開情緒語料庫被實驗於語音情緒辨識的研究上，但是並沒有一個語料庫以台灣常用的語言所錄製。語音情緒辨識會因為語音中不同的文本和語言等資訊影響最後的辨識結果，因此公開的外國語言語料庫不一定適用於台灣語言的情緒辨識研究中。為了解決語料庫的問題我們自行錄製的台灣語言情緒語料庫，採用了最普遍的國語、台語以及客家語三種語言。我們的語料庫仿照德語公開語料庫 EMO-DB 錄製，對每種語言採用十位語者、十句文本以及七種情緒，每個語言收錄七百句，並且在錄製完後進行人工辨識，以一定的人工辨識率做為篩選條件，以確保一定的辨別度。錄製完後以一個龐大的聲學特徵集搭配支持向量機作為後端分類器，以此實驗做為三種語言的基準辨識率。

**關鍵字：**語音情緒辨識，情緒語料庫建置

### 一、緒論

在人類社會中言語溝通扮演著非常重要的角色，人們透過語言來交流和傳遞訊息，基於此理念下發展出許多人機介面科技(human-computer Interaction, HCI)，如蘋果公司的 "Siri" 便是此語音領域的代表作之一，將人類從鍵盤滑鼠輸入限制到能不需手動 (Touchless) 的聲控操作。而人類的交流不只單單透過語言文字，還有情感的交流，情感表達能包含人們所處的狀態，因此有了情感運算的研究。情感運算橫跨許多領域如電腦科學、心理學、認知哲學與工程學等，在語音領域相關為語音情緒辨識及情緒語音合成。基本的情感辨識透過分析人的臉部表情以及聲音來做辨識，更進一步則同時分析對話語意作為分辨依據。

[1]這本小書開始了情感運算的時代，此書定義了情感運算，從一個資訊工程研究者的角度來描述與說明情感運算的應用及重要性，解釋基本的訊號處理與機器學習分類的觀念。情感運算為一個龐大的研究議題，其中橫跨多個領域如資訊科學、心理學、感知科學和工程學等等。[2]對語音情緒辨識做了大略的概論，首先整理了研究中普遍常見的情緒語料庫，探討基本特徵參數如音高(Pitch)、能量(Energy)、基頻(Fundamental Frequency, F0)和梅爾倒頻譜係數(Mel Frequency Cepstral Coefficients, MFCC)等，後端則實驗高斯混合模型(Gaussian Mixture Model, GMM)、支持向量機(Support Vector Machine, SVM)、隱藏式馬可夫模型(Hidden Markov Model, HMM)和類神經網絡(Artificial Neural Networks, ANN)。

研究語音情緒辨識就必須具有情緒語音語料庫用於實驗，而語料庫的收錄會直接影響研究的實驗結果，目前常用的公開情緒語料庫如德語的 Berlin Database of Speech Emotion (EMO-DB)[3]、FAU Aibo[4]和 Vera-Am-Mittag(VAM)[5]。[6]提及在收集情感資料時，人們處於實驗室等非平常習慣的場合所收集到的情感資料實際上會與真實世界表達出的有所差異，因此情緒語音語料庫的錄製大致能分為兩類；第一類為引導性錄音，此類多在實驗室或者是錄音室中錄製，如德語的 EMO-DB，其透過在錄音室錄製並且有三位語言學專家在旁監督與指導，如此可以確保錄製出乾淨並且情緒表達度和差異性高的情緒語料。此種錄製方式較需要語者的表達能力，在語者表達能力較差的情況下無法在錄音室錄製出含有足夠情緒表達的語句；相對的第二種則是自發性的情緒表達語句，如 FAU Aibo 和 VAM，FAU Aibo 透過錄製小孩子與機器人的互動，便可以直接收集到小孩子自然表達的情緒語句。而 VAM 則是從談話性節目中節錄片段再從中挑選出情感較為活躍的句子，如此收錄的語料庫也較自然呈現。

在語音情緒辨識中提出作為基準實驗的有兩者。由於情感辨識相對於語音辨識及語者辨識為較新的領域，因此較早並沒有一個公認的基準可供參考比較，所以 INTERSPEECH 2009 Emotion Challenge[7]便採用了具有明確訓練集及測試集的 FAU Aibo 情緒語料庫，在前端使用過零率(zero crossing rate, ZCR)、能量、基頻、泛音噪音比(harmonics-to-noise, HNR)和梅爾倒頻譜系數。為了解決 FAU Aibo 五類資料量中有兩類佔了近 80%的資料數量嚴重不平均的狀況，使用 SMOTE(Synthetic Minority Oversampling Technology)[8]方法來增加數量較少的類別的樣本數，後端則以隱藏式馬可夫模型以及支持向量機做為辨識器，以此做為 FAU Aibo 的一個實驗基準。[9]實驗 9 個情緒語料庫分別為 Airplane Behaviour Corpus (ABC)、Audiovisual Interest Corpus (AVIC)、Danish Emotional Speech (DES)、EMO-DB、eINTERFACE、Sensitive Artificial Listener (SAL)、SmartKom、Speech Under Simulated and Actual Stress (SUSAS)、和 VAM 等資料庫，這些語料庫分別有不同的錄音條件，包含引導式錄音、自發性語音與各種的錄音環境等等。實驗分為兩種，第一種為擷取一般音框特徵並搭配隱藏式馬可夫模型/高斯混合模型做為後端辨識器，另一種為截取大量的聲學資訊並經由泛函將音框特徵轉換為句子特徵，如此以方便對應至後端支持向量機的訓練與辨識，此篇可視為各情緒語料庫採用聲學模型時的基準實驗參考。

目前常見的公開語音情緒辨識的語料庫並沒有以台灣本土語言錄置而成，而語音中包含了語言、語者、文本、腔調等等的資訊，這些皆會影響語音情緒辨識的準確度。因此為了提供一個較適合用於台灣語言的情緒語料庫，我們自行錄製了台灣最普遍的國語、台語及客家語三種語言的情緒語料庫。我們仿造 EMO-DB 錄製方式錄製，其中相同的七種情緒、五位男性五位女性共十位語者以及十句文本，錄製完後進行人工辨識測驗。最後再採用與[9]相同的基準實驗設定，作為我們語料庫的基準辨識率。

以下為整篇論文的架構，第一節為緒論，第二節詳細介紹我們仿造的德語語料庫 EMO-DB，第三節詳列我們自行錄製的台灣語料庫資訊，第四節為台灣語料庫的基準實驗，最後為我們的結論。

## 二、現有公開語料庫

我們錄製的語料庫參考已廣泛實驗於語音情緒研究的德語語料庫 Berlin Database of Emotion Speech(EMO-DB)。EMO-DB 由 10 位德語母語語者所錄製而成，經由麥克風錄製 48kHz 後降為 16kHz。錄音環境在隔音良好及具有高品質錄音設備的錄音室，因此表達的情緒並非自然呈現。語者由 3 位語言學專家從 40 位新聞播報員中挑選，在面試時每人對每種情緒透過麥克風錄製一句測試句，再將這些句子交予三位語言專家，專家根據自然度及可辨識度挑選出男女各 5 位作為錄音語者。語料庫共採用 7 種情緒分別為中性、生氣、害怕、開心、傷心、噁心和無聊，文本設計為五句長句及五句短句，文本內容為日常生活中的句子並且不帶有任何情緒詞彙，如此使語者在錄音時能較自然的表達語句，且在情緒表達過程中不會受到文本的內容而影響。錄製時語者對相同一句分別以七種情緒表達，因此每位語者至少收錄 70 句，整個語料庫至少包含 700 句，最後全部共錄製約 800 句，其中有部分句子保留二至三個版本。為了確保語句的品質，請 20 位測試者以隨機輪播的方式將每一句做七選一的人工辨識，並且對句子的自然度以 1 至 100 評分，測驗結束後捨棄辨識率低於 80% 和自然度低於 60% 的句子，篩選後七種情緒的句數分別為生氣 127 句、無聊 81 句、噁心 46 句、害怕 69 句、開心 71 句、傷心 62 句和中性 79 句共 535 句。

## 三、台灣情緒語料庫

為了研究跨語言的情緒語音辨識實驗，我們仿照了 EMO-DB 的設計方式錄製了自己的情緒語料庫，分別用相同的七種情緒表達五句長句以及五句短句，每種語言各錄製 700 句，表一列出我們的十句文本，表二與表三分別為台語與客家語的發音文本。我們選擇台灣常見的三種語言-國語、台語及客家語-作為我們的台灣錄製語言，由於客家語有不同腔調，因此我們挑選語者時選擇最為普遍的四縣腔與海陸腔作為我們的錄製腔調。錄音環境為中山大學電資大樓 F5017B 實驗室，錄音麥克風使用 ATH-AT9942，錄音介面卡為 Sound Blaster X-Fi Surround 5.1，取樣率為 16KHz，錄音時語者與麥克風間距約 20 公分。檔案命名格式為 7 碼，分別為第一碼語言 ID、第二三碼語者 ID，第四碼情緒 ID，第五六七碼文本 ID，如 H03At01 音檔則為客家語、語者 03、情緒為生氣和 t01 文本。接著分別對三種語言各找 5 位男性以及 5 位女性作為錄音者。相同的我們在錄製完後也請了 10 位測試者做人工辨識，但我們的測驗中並沒有自然度的評分。我們從中保留辨識率 60% 以上的句子，表四列出三種語言挑選後各情緒所剩餘的句數，圖一表示三種情緒各個人工辨識率的句數，由圖中可看出多數錄製的人工準確度在 60% 以上，所以我們的情緒語料庫具有一定的情緒辨別度。所有的錄音者和測試者皆為大學生或是研究生，並且非語言學或者音樂學系學生。

表一、台灣情緒語料庫 10 句文本

| 文本 ID | 文本                |
|-------|-------------------|
| t01   | 你的早餐放在桌上          |
| t02   | 晚上他有一個約會          |
| t03   | 最近常常睡不飽           |
| t04   | 等下一起去餐廳吃飯         |
| t05   | 過年要買一件新衣服         |
| t06   | 這禮拜放假，要跟朋友一起出去玩   |
| t07   | 昨天早上出門的時候，外套被鉤子勾到 |
| t08   | 今天一整天都沒吃東西好餓喔     |
| t09   | 這班火車人很多，很多人都沒位子坐  |
| t10   | 早上去騎腳踏車，下午散步去買東西  |

表二、台語對應文本

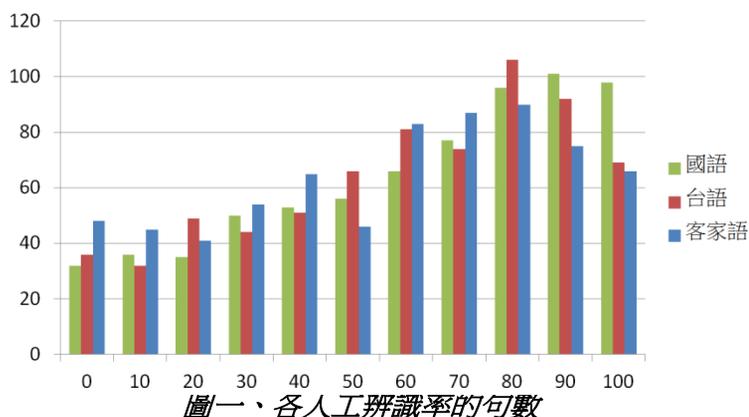
| 文本 ID | 文本                |
|-------|-------------------|
| t01   | 你的早頓放在桌頂          |
| t02   | 暗時依有一個約會          |
| t03   | 最近定定暍不飽           |
| t04   | 等咧做伙去餐廳食飯         |
| t05   | 過年愛買一件新衫          |
| t06   | 這禮拜放假，欲佻朋友做伙出去企桃  |
| t07   | 昨透早欲出門的時陣，外套被鉤仔鉤到 |
| t08   | 今仔日規工攏無食物件就夭      |
| t09   | 這班火車人足濟，真濟人攏無位湯坐  |
| t10   | 透早去騎腳踏車，下晡散步去買物件  |

表三、客家語對應文本

|       |   |
|-------|---|
| 文本 ID | 文本  |
| t01   | (ㄅㄨㄣˊ)(ㄍㄟ)咒九(ㄅㄨㄣˊ)兜揆轟                                       |
| t02   | (ㄅ)補壓(ㄍㄟ)武依炸右(ㄅㄨㄣˊ)   |
| t03   | 追(ㄅㄨㄣˊ)寵寵雖某報  |
| t04   | (den)哈(ㄅㄨㄣˊ)哈(ㄅㄨㄣˊ)蠢堂(ㄅ)翻                                   |
| t05   | 勾輾(ㄅㄨㄣˊ)埋億良心賞夫  |
| t06   | (ㄅㄨㄣˊ)禮辦(ㄅㄨㄣˊ)(ㄅㄨㄣˊ)·(ㄅㄨㄣˊ)老品<br>(ㄅㄨㄣˊ)(ㄅㄨㄣˊ)哈粗(ㄅㄨㄣˊ)(ㄅㄨㄣˊ) |
| t07   | 醜比則損(ㄅㄨㄣˊ)處(ㄅㄨㄣˊ)(ㄅ)死(ㄅㄨㄣˊ)·歪掏(ㄅ<br>ㄨㄣˊ)(ㄅㄨㄣˊ)(ㄅ)(ㄅㄨㄣˊ)豆    |
| t08   | (ㄅㄨㄣˊ)(ㄅㄨㄣˊ)逆(ㄅㄨㄣˊ)逆(ㄅ)宋某奢懂席喊<br>窯喔                         |
| t09   | (ㄅㄨㄣˊ)(ㄅㄨㄣˊ)(ㄅㄨㄣˊ)查擰懂(ㄅㄨㄣˊ)·<br>懂(ㄅㄨㄣˊ)擰朱某(ㄅㄨㄣˊ)(ㄅ)後仇       |
| t10   | 則損(ㄅㄨㄣˊ)(ㄅㄨㄣˊ)師朗查·(ㄅㄨㄣˊ)租三(ㄅㄨㄣˊ)<br>(ㄅㄨㄣˊ)買懂席               |

表四、台灣情緒語料庫各語言篩選後句數

| 情緒 | 國語  | 台語  | 客家語 | 跨語言總數 |
|----|-----|-----|-----|-------|
| 生氣 | 58  | 73  | 62  | 193   |
| 無聊 | 66  | 55  | 59  | 180   |
| 噁心 | 46  | 55  | 54  | 155   |
| 害怕 | 57  | 61  | 64  | 182   |
| 開心 | 67  | 58  | 48  | 173   |
| 傷心 | 57  | 56  | 52  | 165   |
| 中性 | 87  | 64  | 62  | 213   |
| 總數 | 438 | 422 | 401 | 1261  |



圖一、各人工辨識率的句數

## 四、基準實驗

### 4.1 實驗設定

我們實驗中採用 openEAR 0.1.0[10]與 Weka 3.6.9[11]兩套工具用於特徵參數擷取和訓練分類器。輸入的音訊處理採用 25 毫秒的漢明窗(Hamming window)並且處理過程中每次偏移 10 毫秒。在訓練分類器之前使用極值正規法(Min-Max normalization)將特徵參數縮放為 0 至 1，公式定義為算式(1)， $V$  與  $V'$  分別為正規化前與正規化後的值。後端以支持向量機做為分類器，支持向量機的核心函式(kernel function)為基於序列最小化演算法(Sequential Minimal Optimization)的多項式核心(Polynomial Kernel)，Kernel 函式為算式(2)， $p$  設為 1。實驗方式以每次從語料庫中挑出一位語者作為測試語者，而其他語者當作訓練語者(Leave-One-Speaker-Out, LOSO)，經過 10 次訓練與辨識後統計結果便為此語料庫的辨識率。

$$V' = \frac{\text{Max}_V - V}{\text{Max}_V - \text{Min}_V} \quad (1)$$

$$K(x, y) = \langle x, y \rangle^p \quad (2)$$

### 4.2 基準特徵集

本論文使用聲學特徵作為實驗的特徵參數，表五和表六分別列出大型特徵集所包含的低階參數(low-level descriptor, LLD)以及泛函(functional)。首先從音訊中擷取每個音框為 56 維的特徵，再透過泛函將這些音框特徵轉換為一個句子一組的特徵向量，56 個低階參數和 39 個泛函再計算一階與二階動態特徵後共得到 6552 個特徵參數，此便為我們的基準特徵集。

表五、56 個低階參數

| Feature Group | Feature in Group  | # of LLD |
|---------------|---|----------|
| Raw signal    | Zero-crossing-rate  | 1        |
| Signal energy | Logarithm   | 1        |
| Pitch         | Fundamental frequency F0 in Hz via Cepstrum and Autocorrelation (ACF). Exponentially smoothed F0 envelope.  | 2        |
| Voice Quality | Probability of voicing  | 1        |
| Spectral      | Energy in bands 0-250 Hz, 0-650 Hz, 250-650 Hz, 1-4 kHz 25%, 50%, 75%, 90% roll-off point, centroid, flux, and rel. pos. of spectrum max. and min. & 12 Mel-spectrum & Band 1-26 & 26 Cepstral & MFCC 0-12 & 13 | 12       |
| Mel-spectrum  | Band 1-26   | 26       |
| Cepstral      | MFCC 0-12   | 13       |

表六、39 個泛函

| Functionals, etc.   | # of fun. |
|---|-----------|
| Respective rel. position of max./min. value   | 2         |
| Range (max.-min.)   | 1         |
| Max. and min. value - arithmetic mean   | 2         |
| Arithmetic mean, Quadratic mean   | 2         |
| Number of non-zero values   | 1         |
| Geometric, and quadratic mean of non-zero values  | 2         |
| Mean of absolute values, Mean of non-zero abs. values   | 2         |
| Quartiles and inter-quartile ranges   | 6         |
| 95% and 98% percentile  | 2         |
| Std. deviation, variance, kurtosis, skewness  | 4         |
| Centroid  | 1         |
| Zero-crossing rate  | 1         |
| # of peaks, mean dist. btwn. peaks, arth. mean of peaks, arth. Mean of peaks - overall arth. Mean | 4         |
| Linear regression coefficients and corresp. approximation error                                   | 4         |
| Quadratic regression coefficients and corresp. approximation error                                | 5         |

動態特徵計算為 delta regression，係數差公式採用 Hidden Markov Model Tool Kit 定義如算式(3)，其中 W 設為 2。

$$\Delta x_t = \frac{\sum_{i=1}^W (x_{t+i} - x_{t-i}) \cdot i}{2 \sum_{i=1}^W i^2} \quad (3)$$

### 4.3 實驗

根據表五與表六的 6552 維大型特徵集搭配支持向量機作為國語、台語和客家語三種語言的基準實驗，對每個語言的 10 位語者進行 LOSO 實驗後分別得到國語 68.5%、台語 50.7%和客家語 58.5%辨識率。表七至表九分別列出國語、台語和客家語三種語言的詳細辨識結果。

從表七至表九中可看出雖然在人工辨識上可得到至少 60%的辨識率，但在基準實驗中許多情緒的辨識率不及 60%。此實驗結果也反映出錄音狀況，七種情緒中以生氣、無聊和害怕三種情緒的辨識率較佳，此三種情緒對錄音者來說較好表達；而錄音者普遍認為噁心只單純利用語調來表達並不夠直覺，因此在錄音的情緒表達上會相對較差，三種語言的最高辨識率僅 50%；另在錄音過程中傷心和無聊兩種情緒對語者表達來說會有相似的狀況，實驗也可看出傷心的辨識結果有不少部分被分類為無聊，因此造成傷心的辨識率普遍較低。在台語的中性僅 17.2%的準確率，推測因各個語者在表達方式上的差異性過大造成分類器無法順利辨識，主要為語速與聲調的差異，而人工辨識時人們可以由語者的特性去推敲出其情緒，但辨識器上卻無法做出此判斷，所以造成辨識率不理想的狀況。

表七、國語基準實驗混淆矩陣

| 結果 \ 答案 | 生氣 | 無聊 | 噁心 | 害怕 | 開心 | 中性 | 傷心 | 準確率  |
|---------|----|----|----|----|----|----|----|------|
| 生氣      | 44 | 1  | 2  | 0  | 8  | 0  | 3  | 75.9 |
| 無聊      | 0  | 59 | 0  | 0  | 0  | 4  | 3  | 89.4 |
| 噁心      | 6  | 4  | 21 | 0  | 7  | 4  | 4  | 45.7 |
| 害怕      | 2  | 0  | 0  | 45 | 6  | 2  | 2  | 78.9 |
| 開心      | 7  | 0  | 4  | 6  | 40 | 1  | 9  | 59.7 |
| 傷心      | 1  | 10 | 7  | 2  | 3  | 31 | 3  | 54.4 |
| 中性      | 5  | 4  | 4  | 3  | 10 | 1  | 60 | 69.0 |

表八、台語基準實驗混淆矩陣

| 結果 \ 答案 | 生氣 | 無聊 | 噁心 | 害怕 | 開心 | 中性 | 傷心 | 準確率  |
|---------|----|----|----|----|----|----|----|------|
| 生氣      | 45 | 1  | 8  | 0  | 14 | 0  | 5  | 61.6 |
| 無聊      | 0  | 43 | 0  | 0  | 0  | 7  | 5  | 78.2 |
| 噁心      | 6  | 4  | 27 | 2  | 4  | 8  | 4  | 49.1 |
| 害怕      | 7  | 1  | 4  | 39 | 6  | 2  | 2  | 63.9 |
| 開心      | 12 | 5  | 4  | 3  | 24 | 1  | 9  | 41.4 |
| 傷心      | 0  | 16 | 8  | 3  | 1  | 25 | 3  | 44.6 |
| 中性      | 16 | 13 | 6  | 6  | 10 | 2  | 11 | 17.2 |

表九、客家語基準實驗混淆矩陣

| 結果 \ 答案 | 生氣 | 無聊 | 噁心 | 害怕 | 開心 | 中性 | 傷心 | 準確率  |
|---------|----|----|----|----|----|----|----|------|
| 生氣      | 44 | 1  | 2  | 0  | 6  | 1  | 4  | 72.6 |
| 無聊      | 0  | 59 | 0  | 0  | 0  | 6  | 1  | 86.7 |
| 噁心      | 6  | 4  | 21 | 0  | 2  | 3  | 5  | 50.0 |
| 害怕      | 2  | 0  | 0  | 45 | 6  | 3  | 6  | 64.1 |
| 開心      | 7  | 0  | 4  | 6  | 16 | 2  | 8  | 33.3 |
| 傷心      | 1  | 10 | 7  | 2  | 1  | 15 | 2  | 28.8 |
| 中性      | 5  | 4  | 4  | 3  | 5  | 0  | 39 | 62.9 |

## 五、結論

在此篇論文中我們建立了一個台灣語言的情緒語料庫，其中以台灣常見的國語、台語和客家語三種語言錄製而成。我們仿照公開語料庫 EMO-DB 錄製，共包含了生氣、無聊、噁心、害怕、開心、中性和傷心共七種，每種語言的語者為十位，分別為五位男性以及五位女性，文本共十句，每位語者以七種情緒來表達同一句文本。在錄製後進行人工辨識的檢驗，以人工辨識率 60% 作為篩選基準，以此確保我們語料具有一定的辨別度，篩選後三種語言的所剩句數為國語 438 句、台語 422 句和客家語 401 句。另我們使用了一個龐大的聲學特徵集並搭配支持向量機做為我們的基準實驗，其中分別在三種語言可以得到國語 68.5%、台語 50.7% 和客家語 58.5% 的辨識率，便為三個語言的基準辨識率。此資料庫及實驗數據可做為未來台灣語音情緒辨識使用及參考。在未來若可以我們希望能夠更提升語料庫的品質，如改善台語的中性辨識率過低的狀況，對於篩選後句數剩餘過少的語者進行重錄或以新的語者資料代替，提升品質的同時也增加語料庫的總句數。

## 參考文獻

- [1] R. W. Picard, *Affective computing*. MIT Press, 1997.
- [2] M. El Ayadi, M. S. Kamel, and F. Karray, “Survey on speech emotion recognition: features, classification schemes, and databases,” in *Pattern Recognition*, vol. 44, pp. 572–587, Mar. 2011.
- [3] F. Burkhardt, A. Paeschke, M. Rolfes, W. F. Sendlmeier, and B. Weiss, “A database of german emotional speech,” in *Proceedings of INTERSPEECH*, pp. 1517–1520, 2005.
- [4] S. Steidl, *Automatic classification of emotion related user states in spontaneous children’s speech*. PhD thesis, University of Erlangen-Nuremberg, 2009.
- [5] M. Grimm, K. Kroschel, and S. Narayanan, “The Vera am Mittag german audio-visual emotional speech database,” in *Proceedings of IEEE International Conference on Multimedia and Expo*, pp. 865–868, 2008.
- [6] D. McDuff, R. Kaliouby, and R. Picard, “Crowdsourcing facial responses to online videos,” in *IEEE Transactions on Affective Computing*, vol. 3, pp. 456–468, 2012.
- [7] B. Schuller, S. Steidl, and A. Batliner, “The INTERSPEECH 2009 emotion challenge,” in *Proceedings of INTERSPEECH*, pp. 312–315, 2009.
- [8] G. Weiss and F. Provost, “The effect of class distribution on classifier learning: An empirical study,” tech. rep., Department of Computer Science, Rutgers University, 2001.
- [9] B. Schuller, B. Vlasenko, F. Eyben, G. Rigoll, and A. Wendemuth, “Acoustic emotion recognition: A benchmark comparison of performances,” in *Proceedings of IEEE Workshop on of Automatic Speech Recognition Understanding (ASRU)*, pp. 552–557, 2009.
- [10] F. Eyben, M. Wollmer, and B. Schuller, “OpenEAR introducing the munich open-source emotion and affect recognition toolkit,” in *Proceedings of ACII International Conference on Affective Computing and Intelligent Interaction and Workshops*, 2009, pp. 1–6.
- [11] I. H. Witten, E. Frank, L. Trigg, M. Hall, G. Holmes, and S. J. Cunningham, “Weka: Practical machine learning tools and techniques with java implementations,” 1999.