

# Keystroke Dynamics

Laurențiu-Iulian Iordache-Stoicescu

**Abstract**—În ziua de astăzi, calculatoarele sunt folosite peste tot pentru a stoca și procesa o gamă largă de date. De asemenea, numărul atacurilor cibernetice a crescut și el. Pentru a putea proteja aceste sisteme de intruși, utilizarea unui sistem de securitate adecvat reprezintă o prioritate. În prezent sunt utilizate cu succes mai multe metode de securitate bazate pe măsuri biometrice, cum ar fi analiza amprentelor sau recunoașterea facială. Această cercetare se va axa însă pe o metodă prea puțin folită și anume, dinamica apăsării tastelor, cunoscută în literatură drept „Keystroke dynamics”.

## Index Terms—

## I. INTRODUCERE

### A. Necesitate

Calculatoarele au devenit omniprezente în societatea modernă, conform statisticilor oferite de Internet World Stats [1], numărul utilizatorilor unici de internet calculat la sfârșitul lunii Iunie 2018 este de aproximativ 4.2 miliarde. Pe cât este de mare numărul de utilizatori ce au acces la internet, pe atât de mare este numărul de atacuri cibernetice ce îi vizează. Conform datelor oferite de ..., aproximativ 63% din toate intruziunile în rețele și furturile de informații se datorează compromiterii datelor de autentificare. Un atac cibernetic cunoscut care a constat în furtul datelor personale ale aproximativ 500 de milioane de conturi este cel ce a vizat site-ul „yahoo.com”.

Din moment ce depindem din ce în ce mai mult de calculatoare, iar riscurile folosirii acestora cresc de la o zi la alta, este normal ca și nivelul de securitate să fie sporit pentru a putea face față atacurilor. Utilizarea de metrice biometrice în procesul de autentificare este unul dintre pașii făcuți pentru sporirea securității în ceea ce privește autentificare unui utilizator. În acest caz se merge pe ideea că un atacator poate fura identitatea digitală a unui utilizator (utilizator, parola, token etc.) dar nu poate fura sau replica ceea ce este utilizatorul.

În prezent sunt implementate cu succes mai multe metode de autentificare pe baza de metrice biometrice cum ar fi recunoașterea utilizatorului pe bază de amprentă papilară, recunoașterea facială sau a irisului. Aceste metode însă necesită componente hardware suplimentare pentru achiziția datelor biometrice. O altă metodă de identificare a unui utilizator, mai puțin populară, o reprezintă analiza dinamicii apăsării tastelor. Această metodă are avantajul că nu necesită componente hardware adiționale, deoarece orice calculator are o tastatură. Pe lângă acesta mai are un avantaj semnificativ și anume că poate realiza achiziția metricilor în timp ce utilizatorul își îndeplinește sarcinile uzuale fără a-l deranja și fără a se face sesizată achiziția de date, aceasta fiind o metodă neintruzivă.

### B. Generalități biometrie

### C. Dinamica apăsării tastelor

Măsura biometrică a tastării este referită în literatură sub forma de „Keystroke dynamics” (dinamica apăsării tastelor). Dinamica apăsării tastelor se referă la felul în care o persoană apasă tastele unei tastaturi. Această metodă este bazată pe caracteristicile de scriere ale persoanelor cum ar fi durata apăsării unei taste, latența dintre apăsări consecutive ale tastelor, timpul dintre două apăsări consecutive și în cazul în care se poate, forța apăsării. Cele mai folosite metrice

sunt: timpul de apăsare, acesta reprezintă durata de timp a menținerii unei taste apăsată și timpul de pauză, care reprezintă durata de timp dintre eliberarea unei chei și apăsarea alteia.

1) *Istorie*: Această metodă este derivată din ideea de indentificare a expeditorului unui cod Morse ce folosește un telegraf. Această tehnică a fost analizată în timpul celui de-al doilea război mondial și poartă numele de „fist of the sender” (pumnul expeditorului). Aceasta a fost utilizată cu succes pentru a monitoriza deplasarea trupelor pe baza recunoașterii tiparului de transmisie a expeditorului mesajului [2].

2) *Mecanismul psihologic*:

## II. ACHIZIȚIA DATELOR

## III. ABORDĂRI ȘI REZULTATE

### A. Baza de Date

O chestiune importantă de luat în calcul în evaluarea unui sistem de recunoaștere a emoțiilor din vorbire este gradul de naturalețe al bazei de date utilizată în evaluarea performanțelor. Se pot trage concluzii incorecte despre sistem dacă datele utilizate nu sunt de calitate.

Din păcate, majoritatea bazelor de date cu date audio ce conțin înregistrări preluate de la vorbitori cu anumite emoții nu sunt disponibile public. Acest fapt determină o îngreunare a procesului de evoluție în această ramură. Cercetătorii nu sunt capabili să se coordoneze din punctul de vedere al greșelilor și le repetă pentru baze de date diferite.

### B. Fapt

În carlinga avioanelor s-a descoperit că sistemele de recunoaștere a vorbirii antrenate cu mesaje de la un subiect aflat sub stres au performanțe mult mai bune decât sistemele antrenate cu mesaje de la subiecți aflați în stare normală.

## IV. EMOȚIILE

Cuvântul emoție este adesea considerat ca fiind o stare afectivă a minții. În conversațiile umane normale, oamenii sunt capabili să identifice emoțiile altora pe baza vocii, a posturii, a gesturilor și a expresiei faciale. Emoțiile sunt adesea neglijate în interacțiunea dintre om și mașină. Pentru a putea detecta emoțiilor trebuie determinate trei aspecte:

- Ce reprezintă o stare afectivă?
- Ce semnale din comunicația verbală exprimă informația despre starea afectivă?
- Cum ar trebui combinate diversele caracteristici pentru a optimiza percepția emoțiilor de către mașină?

În literatură, emoțiile enumerate de Ekman (furie, dezgust, frică, fericire, tristețe și surprindere) sunt folosite cu precădere. Au fost propuse mai multe modele de ilustrare a relațiilor dintre valențe, excitație și emoții. Figura 1 prezintă modelul PAD al stărilor emoționale o analiză a emoțiilor propuse de alt psiholog, Russell [5].

Din acest model se disting următoarele:

- Excitarea (Arousal) - emoțiile sunt caracterizate pe baza nivelului lor de excitație.
- Valență - emoțiile sunt caracterizate în funcție de cât sunt de pozitive sau de negative. Emoțiile pozitive sporesc atenția și alte funcții cognitive ale unei persoane

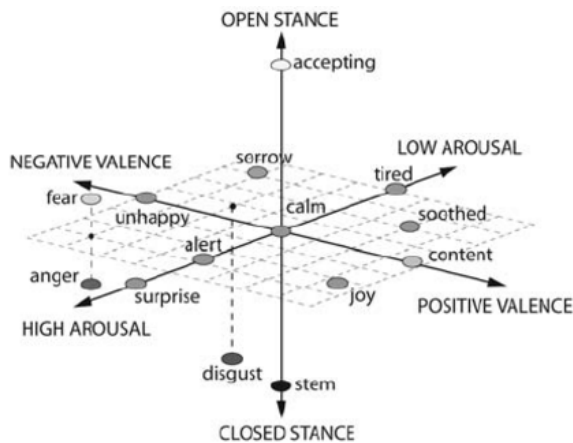


Fig. 1. Modelul PAD 3D de clasificare a emoțiilor [7]

## V. CARACTERISTICI UTILIZATE PENTRU RECUNOAȘTEREA EMOȚIILOR

Pentru realizarea unui sistem de recunoaștere a emoțiilor din vorbirea unui subiect este necesară identificarea caracteristicilor ce caracterizează diversele emoții. O alegerea caracteristicilor va influența performanțele sistemului.

Pentru o analiză eficientă a caracteristicilor trebuie să se țină seama de regiunea de analiză utilizată pentru extragerea caracteristicilor, să se aleagă caracteristica potrivită pentru identificarea emoției, să se ia în calcul efectele proceselor uzuale de procesare a semnalului vocal cum ar fi filtrarea și eliminarea zonelor de liniște dar și de asemenea să se ia în calcul faptul că trăsăturile acustice s-ar putea să nu fie suficiente pentru modelarea emoțiilor și că ar putea să se folosească și alte tipuri de caracteristici cum ar fi informația discursului sau trăsăturile faciale. În continuare vor fi detaliate aceste aspecte.

### A. Caracteristici locale și caracteristici globale

Primul aspect se referă la selectarea regiunii de analiză folosită pentru extragerea caracteristicilor. Aici există două abordări. Divizarea semnalului în cadre de durată scurtă pentru care se calculează vectorul caracteristic sau extragerea de mărimi statistice globale pentru întreaga propoziție.

Deoarece semnalul vocal este nestăionar, se practică divizarea acestuia în ferestre mici de 20 până la 40 de ms. O caracteristică importantă a semnalului vocal, care facilitează analiza acestuia este că acesta este aproximativ staționar pe perioade scurte de timp. Din fiecare astfel de fereastră sunt extrase caracteristici prozodice cum ar fi frecvența fundamentală și energia și denumite caracteristici locale.

Pe de altă parte, caracteristicile globale reprezintă statistici ale tuturor caracteristicilor locale extrase din mesaj. Majoritatea cercetătorilor au agreat faptul că aceste caracteristici globale sunt superioare caracteristicilor locale din punct de vedere al preciziei și a timpului de clasificare. Caracteristicile globale mai au încă un avantaj față de cele locale și anume acela că sunt mult mai puține la număr, ceea ce facilitează o execuție mult mai rapidă a unor algoritmi cum ar fi cel de "validare în cruce". Algoritmul de validare în cruce reprezintă o tehnică de evaluare a modelelor predictive petiționând mostra originală într-un set de antrenare pentru antrenarea modelului și un set de evaluare [2]. Utilizarea acestora este eficientă doar în cazul distingerii emoțiilor de excitație înaltă (ex. furia, frica și bucuria) de cele de excitație joasă (ex. tristețea). Caracteristicile globale nu reușesc să realizeze o clasificare precisă a emoțiilor cu excitație similară de exemplu identificarea furiei față de bucurie. De asemenea, mai

prezintă și dezavantajul că informația temporală prezentă în mesaj este pierdută. Și încă un dezavantaj îl reprezintă chiar faptul că sunt într-un număr mai mic decât caracteristicile locale deoarece s-ar putea să nu se poată antrena modele de tipul HMM sau SVM.

O altă abordare pentru extragerea de caracteristici o reprezintă segmentarea semnalului pe durate scurte pentru analiza fonemelor. Această abordare se bazează pe faptul că forma spectrului fonemelor variază în funcție de emoții. Această afirmație este valabilă doar pentru vocale.

### B. Clasificarea caracteristicilor

Un aspect important în recunoașterea emoțiilor îl reprezintă extragerea caracteristicilor semnalului vocal care caracterizează eficient emoția din mesaj dar în același timp este independentă de vorbitor sau de conținutul lexical.

Caracteristicile vorbirii pot fi grupate în următoarele categorii: Caracteristici continue, caracteristici calitative, caracteristici spectrale și caracteristici bazate pe operatorul TEO. În figura 1 se pot observa exemple de caracteristici aparținând fiecărei categorii.

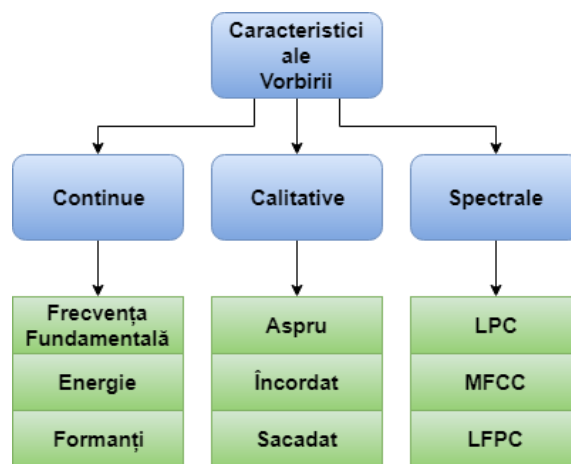


Fig. 2. Categorii de caracteristici ale semnalelor vocale

**1) Caracteristici de vorbire continue:** Caracteristicile prozodice continue cum ar fi frecvența fundamentală și energia exprimă o mare parte din sentimentele existente într-un mesaj. S-a observat că starea de excitație a subiectului influențează distribuția spectrală a energiei precum și frecvența și durata pauzelor din vorbire. Conform studiilor, aceste caracteristici acustice pot fi grupate în [1]:

- Caracteristici asociate frecvenței fundamentale;
- Caracteristici bazate pe formanți
- Caracteristici asociate energiei semnalului vocal
- Caracteristici ale articulării

**2) Caracteristici de calitate a vocii:** Calitatea vocii este strâns legată de tipurile de emoții, aceasta este descrisă de emoții care direcționează puternic subiecții către un șir de acțiuni. Acestea sunt opuse emoțiilor fundamentale care influențează pozitiv sau negativ acțiunile și gândurile unei persoane. O gamă largă de variabile fonetice contribuie la impresia subiectivă de calitate a vocii. Corelațiile acustice sunt grupate în următoarele categorii [1]:

- Nivelul vocii, amplitudinea semnalului, energia și durata s-au dovedit a fi metrice bune pentru nivelul vocii.
- Frecvența fundamentală a vocii
- Frază, foneme, cuvinte
- Structurile temporale

3) *Caracteristici bazate pe spectru:* Acestea sunt adesea reprezentări ale semnalului pe perioade scurte de timp. Emoțiile dintr-un mesaj influențează distribuția spectrală de energie a mesajului. De exemplu, mesajele care sunt rostite când subiectul este fericit au un nivel ridicat al energiei pentru frecvențele înalte pe când cele rostite de subiecți care sunt triști emoțional au nivele scăzute ale energiei pentru aceleași frecvențe. Pentru o mai bună exploatare a gamei de frecvențe, spectrul este trecut printr-o serie de filtre trece-bandă. Caracteristicile spectrale sunt extrase pe baza rezultatelor obținute de la fiecare filtru. Deoarece percepția umană a frecvenței fundamentale nu este modelată liniar, filtrele trece-bandă sunt distribuite uniform cu privire la o metodă neliniară potrivită cum ar fi scala frecvențelor Mel.

### C. Procesarea vorbirii

Înainte extragerii caracteristicilor este necesară o preprocesare a semnalului audio. De exemplu, datorită diferențelor din mediul de înregistrare este necesară o normalizare a energiei pentru toate mesajele. Alt exemplu îl reprezintă netezirea contururilor extrase, pentru aceasta se folosește metoda suprapunerii cadrelor. Pentru eliminarea ondulațiilor din spectrul de frecvențe sunt utilizate ferestre de tip Hamming.

Deoarece intervalele de liniște pot oferi informații importante asupra stării emoționale, acestea sunt păstrate intacte, față de procesele normale de analiză a vorbirii.

Odată extrase caracteristicile se poate să fie necesară o postprocesare înainte de a antrena clasificatorul. De exemplu, este posibil ca vectorii extrași să aibă unități diferite și prin urmare, valorile lor numerice pot avea ordine de mărime diferite sau pot lipsi cu desăvârșire.

### D. Combinarea caracteristicilor acustice cu alte surse de informații

În multe situații, nu este suficientă doar informația provenită din semnalul vocal. Se pot utiliza informațiile referitoare de expresiile faciale sau informația mesajului pentru a putea îmbunătăți performanțele de recunoaștere

1) *Utilizarea informațiilor acustice și lingvistice:* Semnificația mesajului vorbit este un bun indicator al emoției transmise. Pentru a putea utiliza informațiile lingvistice este necesară recunoașterea secvenței de cuvinte din propoziție. Pentru aceasta este necesar un model de limbă. Acesta descrie constrângerile posibilelor secvențe de cuvinte ale unei limbi. Un astfel de model de limbă este modelul N-gram, acesta atribuie probabilități secvențelor de cuvinte ce pot avea loc. În figura 2 este prezentată arhitectura de bază a unui sistem de recunoaștere a vorbirii ce combină rolurile modelului acustic cu cel al modelului lingvistic pentru identificarea celei mai bune secvențe în care am putea detecta o anumită emoție.

Modelul de limbă este realizat pe baza transcrierii cuvintelor, în paralel are loc extragerea de vectori de caracteristici din semnalul vocal. Vectorul rezultat împreună cu dicționarul de limbă sunt utilizate pentru a antrena modelele acustice de foneme. La recunoaștere sunt utilizate atât modelul acustic cât și cel de limbă pentru a putea recunoaște secvența după următoarea regulă:

$$\begin{aligned}\hat{W} &= \arg \max_w P(W|Y) = \arg \max_w \frac{P(W)P(Y|W)}{P(Y)} \\ &= \arg \max_w P(W)P(Y|W)\end{aligned}\quad (1)$$

Formula utilizează funcția argmax, care selectează argumentul ce maximizează probabilitatea secvenței de cuvinte. Dezvoltarea din ecuația 1 are la bază regula Bayes și s-a făcut ținând cont de faptul

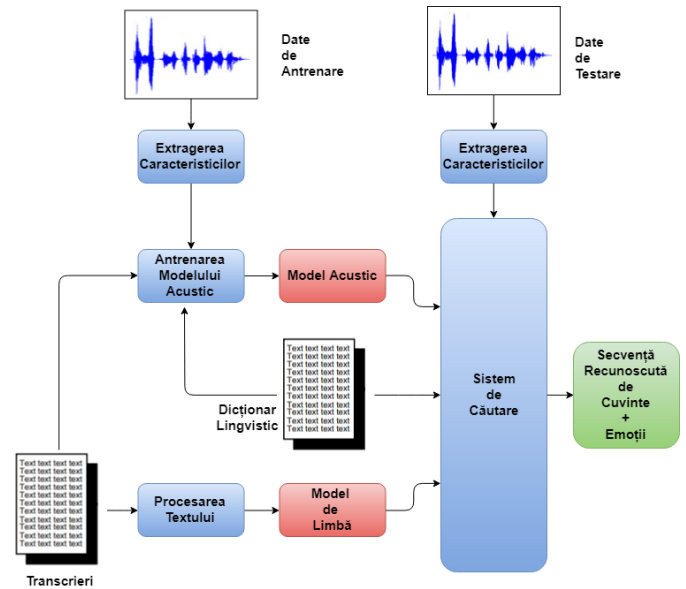


Fig. 3. Arhitectura unui sistem de recunoaștere a emoțiilor ce utilizează informația acustică și lingvistică

că probabilitatea mesajului vorbit  $p(X)$  este independentă de secvența de cuvinte  $W$ . Ultimul rezultat evidențiază doi factori care pot fi estimați direct. Problema inițială (găsirea secvenței de cuvinte pe baza mesajului vorbit) a fost împărțită în două sub-probleme mai simple: a) estimarea probabilității apriori a secvenței de cuvinte  $p(W)$  și b) estimarea probabilității mesajului vorbit dată fiind secvența de cuvinte pronunțată  $p(X|W)$ . Primul factor poate fi estimat utilizând exclusiv un model de limbă, iar cel de-al doilea poate fi estimat cu ajutorul unui model acustic. Cele două modele pot fi construite independent așa cum se va vedea în secțiunea următoare, dar vor fi folosite împreună pentru a decoda un mesaj vorbit, așa cum arată ecuația 1 [3].

### 2) Utilizarea informațiilor acustice, lingvistice și de discurs:

Indicatorii de discurs reprezintă expresii lingvistice care transmit informații explicite despre structura discursului sau au o contribuție semantică la aceasta. În contextul de recunoaștere a emoțiilor, informațiile legate de discurs mai prezintă și modul în care utilizatorul interacționează cu sistemul. Se întâmplă adesea ca utilizatorul să manifeste emoții în timpul interacțiunii cu sistemul, emoții cum ar fi frustrarea. Informațiile de discurs sunt combinate cu informațiile acustice pentru a îmbunătăți performanțele de recunoaștere.

## VI. METODE DE RECUNOAȘTERE A EMOȚIILOR

În continuare se va prezenta o comparație între rezultatele obținute de două metode de recunoaștere a emoțiilor. Prima metodă se bazează pe utilizarea modele de mixturi Gaussiene antrenate pe baza caracteristicilor globale cum ar fi frecvența fundamentală și conturul energiei semnalului audio. A doua metodă utilizează modele Markov ascunse antrenate pe baza caracteristicilor locale ale semnalului audio. De asemenea, vor mai fi prezentați și alți algoritmi utilizați în detecția emoțiilor împreună cu rezultatele obținute.

### A. Modele de mixturi Gaussiene

Modelele de mixturi Gaussiene oferă o bună aproximare a caracteristicilor extrase din semnalul audio prin amestecul ponderat de densități Gaussiene. Coeficienții mixturilor s-au calculat cu ajutorul unui algoritm de maximizare a așteptării. Fiecare emoție este

Labeled emotion	Recognized emotion						
	ang	dis	fea	sur	joy	ntl	sad
ang	<b>91.1</b>	1.2	0.6	6.3	0.4	0.1	0.3
dis	5.4	<b>76.8</b>	6.7	0.1	6.8	3.2	1.0
fea	0.2	6.4	<b>82.8</b>	0.6	3.0	0.3	6.7
sur	2.4	2.2	3.0	<b>87.2</b>	4.6	0.1	0.5
joy	3.0	0.7	0.8	0.0	<b>93.2</b>	0.2	2.1
ntl	0.2	3.4	0.4	0.5	2.7	<b>89.6</b>	3.2
sad	0.2	0.1	5.8	3.8	0.4	2.2	<b>86.6</b>

Fig. 4. Matricea de confuzie a analizei utilizând modele GMM [8]

modelata într-un GMM. În experiment s-a remarcat un rezultat bun utilizând un număr de 16 mixturi Gaussiene.

În tabel sunt următoarele abrevieri: sur - surprise, joy - joy, ang - anger, fea - fear, dis - disgust, sad - sadness și ntl - neutral [8].

### B. Modele Markov ascunse

Modelul HMM reprezintă un automat cu stări finite alcătuit dintr-un set de stări conectate între care se tranzizionează. Secvența de stări a modelului HMM este ascunsă, cu toate Matricea de confuzie a fost realizată utilizând 64 de stări și patru mixturi Gaussiene.

Labeled emotion	Recognized emotion						
	ang	dis	fea	sur	joy	ntl	sad
ang	<b>68.5</b>	12.7	2.6	1.8	2.7	8.4	3.3
dis	12.8	<b>84.7</b>	2.1	0.3	0.0	0.1	0.0
fea	1.8	0.1	<b>95.4</b>	0.2	2.0	0.4	0.1
sur	6.3	6.7	6.3	<b>73.5</b>	6.1	0.9	0.2
joy	10.1	11.8	7.9	1.2	<b>68.0</b>	0.5	0.5
ntl	10.4	0.9	1.0	0.1	1.9	<b>79.6</b>	6.1
sad	5.9	10.1	2.8	2.1	2.2	1.8	<b>75.1</b>

Fig. 5. Matricea de confuzie a analizei utilizând modele HMM [8]

În cazul utilizării de modele Markov ascunse s-a obținut o acuratețe medie de 77.8%.

În figura 6 este prezentată evoluția ratei de recunoaștere în funcție de numărul de stări folosite. S-a remarcat că cele mai bune rezultate se obțin pentru utilizarea a 64 de stări.

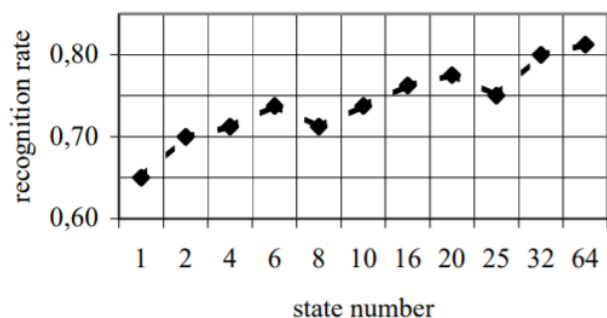


Fig. 6. Rata de recunoaștere în funcție de numărul de stări utilizate [8]

După cum se poate observa, o creștere generală a performanțelor poate fi obținută crescând numărul de stări. Creșterea performanței vine însă cu un cost și anume creșterea timpului de procesare.

### C. K-nearest neighbors

În procesul de recunoaștere de tipare, k-nearest neighbors reprezintă un algoritm non-parametric folosit pentru clasificare și regresie. Acesta primește ca intrare un număr K de caracteristici de antrenare pe baza cărora poate determina o separare între emoții. În acest caz s-au utilizat 70% din setul de date s70 pentru antrenare și 30% dintre ele ca set de test. Pentru K varian între 1 și 15 s-au obținut următoarele rezultate de precizie: Pentru utilizarea a 8 caracteristici s-a obținut o acuratețe medie aproximativ egală cu 55%, în cazul utilizării a 10 și 14 caracteristici se obține un rezultat mult mai mare al detecției de furie și anume 65%. Se obțin rezultate slabe în cazul recunoașterii fricii (13%, 7% și 1% pentru un număr de 8, 10 respectiv 14 caracteristici).

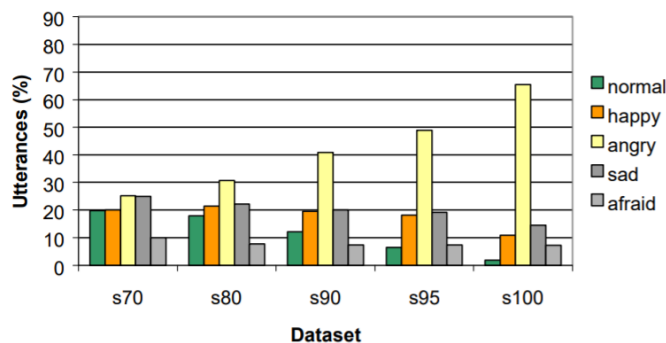


Fig. 7. Distribuția emoțiilor pentru setul de date [9]

### D. Rețele neuronale

S-a utilizat o reșea cu două niveluri cu un vector de intrare de 8, 10 respectiv 14 elemente, 10 sau 20 de noduri în nivelul ascuns și 5 noduri în nivelul liniar de ieșire. Numărul de intrări corespunde numărului de caracteristici iar numărul de ieșiri corespunde numărului de emoții. Pentru antrenarea acestei rețele s-au utilizat seturile de date s70, s80 și s90 care au fost împărțite în 70% date de antrenare și 30% date de testare. Au fost create mai multe clasificatoare ce au fost antrenate cu matrici de ponderare diferite. Această abordare a fost aplicată pe setul de date s70 pentru 8 caracteristici. S-au obținut următoarele rezultate:

TABLE I

REZULTATE OBTINUTE ÎN CAZUL RECUNOAȘTERII DE EMOȚII CU AJUTORUL REȚELOR NEURONALE

Acuratețe	Emoție
Stare normală	55-65%
Fericire	60-70%
Furie	60-80%
Frică	25-50%
Acuratețe medie	65%

### E. Ansamblu de rețele neuronale

Un ansamblu de rețele neuronale este compus dintr-un număr impar de clasificatoare ce au fost antrenate pe diferite subseturi de date folosind agregare bootstrap dar și tehnici de validare în cruce. Deciziile se iau pe baza principiului de vot majoritar. Numărul de rețele neuronale din ansamblu a fost variat între 7 și 15. În figura 8 este prezentată acuratețea de recunoaștere medie a ansamblului format din 15 rețele neuronale. S-a utilizat setul de date s70, numărul de caracteristici utilizate este de 8, 10 respectiv 14 și s-a ținut cont și de cele două arhitecturi cu 10 respectiv 20 de neuroni în nivelul ascuns.

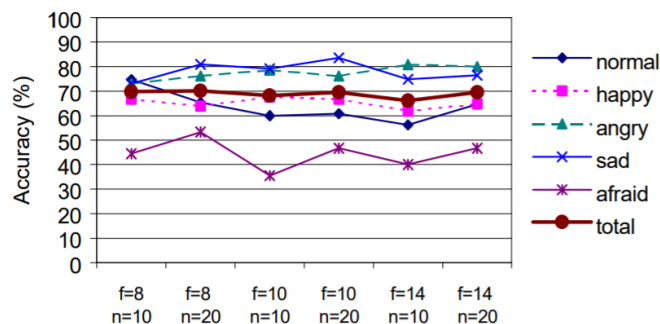


Fig. 8. Acuratețea recunoașterii emoțiilor pentru setul de date s70 [9]

Se poate observa din grafic că acuratețea cu care se efectuează detecția fericirii își păstrează valoarea pe la aproximativ 65% pentru toate cazurile de variere a caracteristicilor de intrare sau a arhitecturilor. Acuratețea pentru detecția fricii este relativ scăzută, aceasta este identificată cu o precizie de 35-53%. Precizia detecției furiei pornește de la 73% pentru un vector de 8 caracteristici și crește până la 81% pentru un set de 14 caracteristici. Cel mai bun rezultat din punct de vedere al acurateții îl reprezintă detecția tristeții, aceasta având o precizie de 73% to 83%. Acuratețea medie este situată la 70%.

## VII. CONCLUZII

În acest document au fost prezentate câteva aplicații practice în care poate fi utilizat procesul de recunoaștere a emoțiilor pe baza semnalului vocal. De asemenea au mai fost prezentate și principalele caracteristici pe baza cărora se face analiza semnalului și s-au prezentat rezultatele obținute în cazul utilizării a două metode distincte de recunoaștere a emoțiilor. Ambele metode prezentate s-au dovedit a fi capabile să identifice emoțiile din mesaje audio vocale. Tabelele de confuzie arată clar că unele emoții sunt confundate între ele, ceea ce indică similitudini între caracteristicile emoțiilor. De asemenea, aceste rezultate din tabelul de confuzie mai pot fi influențate și de faptul că datele utilizate la antrenare nu au avut cea mai bună calitate din punctul de vedere al calității emoțiilor. Din cele două tabele se remarcă faptul că utilizarea caracteristicilor globale oferă rezultate mai bune decât utilizarea caracteristicilor locale.

## REFERINTE

- [1] <https://www.internetworldstats.com/stats.htm>
- [2] L. F. Coppenrath and Associates. „Biopassword Technology Overview”, <http://www.lfca.net/Reference%20Documents/Biometric%20Solutions%20By%20Classification.pdf>
- [3] Horia Cucu: “Proiect de cercetare-dezvoltare în Tehnologia Vorbirii”
- [4] Senaka Amarakeerthi, Rasika Ranaweera, and Michael Cohen: “Speech-based Emotion Characterization using Postures and Gestures in CVEs”, Cyberworlds (CW), 2010 International Conference, 20-22 Oct. 2010.
- [5] Vicki R. LeBlanc, Meghan M. McConnell, Sandra D. Monteiro: “Predictable chaos: a review of the effects of emotions on attention, memory and decision making”, Adv Health Sci Educ Theory Pract. 2015 Mar; 20(1):265-82.
- [6] Paul Ekman: “Emotions Revealed”
- [7] Jonghwa Kim, Elisabeth Andre: “Emotion Recognition Based on Physiological Changes in Music Listening”, IEEE Transactions on Pattern Analysis and Machine Intelligence ( Volume: 30, Issue: 12, Dec. 2008 ), 02 February 2008, 2067 - 2083
- [8] Björn Schuller, Gerhard Rigoll, and Manfred Lang: “HIDDEN MARKOV MODEL-BASED SPEECH EMOTION RECOGNITION”, Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference, 6-9 July 2003.
- [9] Valery A. Petrushin: “EMOTION RECOGNITION IN SPEECH SIGNAL: EXPERIMENTAL STUDY, DEVELOPMENT, AND APPLICATION”, International Conference on Spoken Language Processing (ICSLP 2000)