

SOLUȚIE MACHINE LEARNING PENTRU RECUȚNOAȘTEREA EMOȚIEI ÎN VORBIRE

Coordonator

Prof. dr. Ing. Ștefan HOLBAN

Candidat

Steleac Raul-Dacian

IUNIE 2020

Recunoașterea emoției în vorbire



Sursa [1]

Om - Om

Informație lingvistică
Vs
Informație emoțională

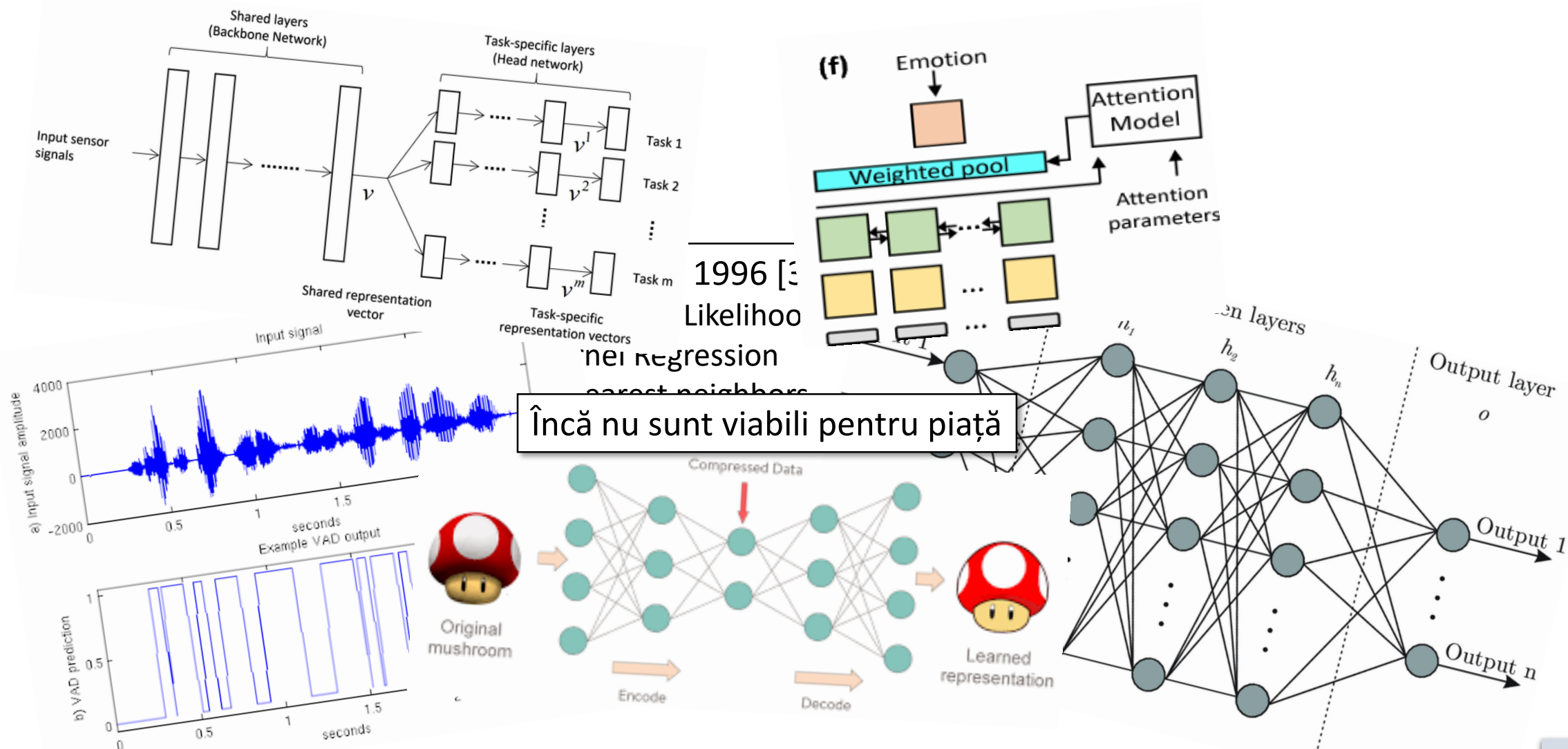


Sursa [2]

Om - Mașină

Noua generație de vorbitori

Domeniul recunoașterii emoției în vorbire



Bazele de date

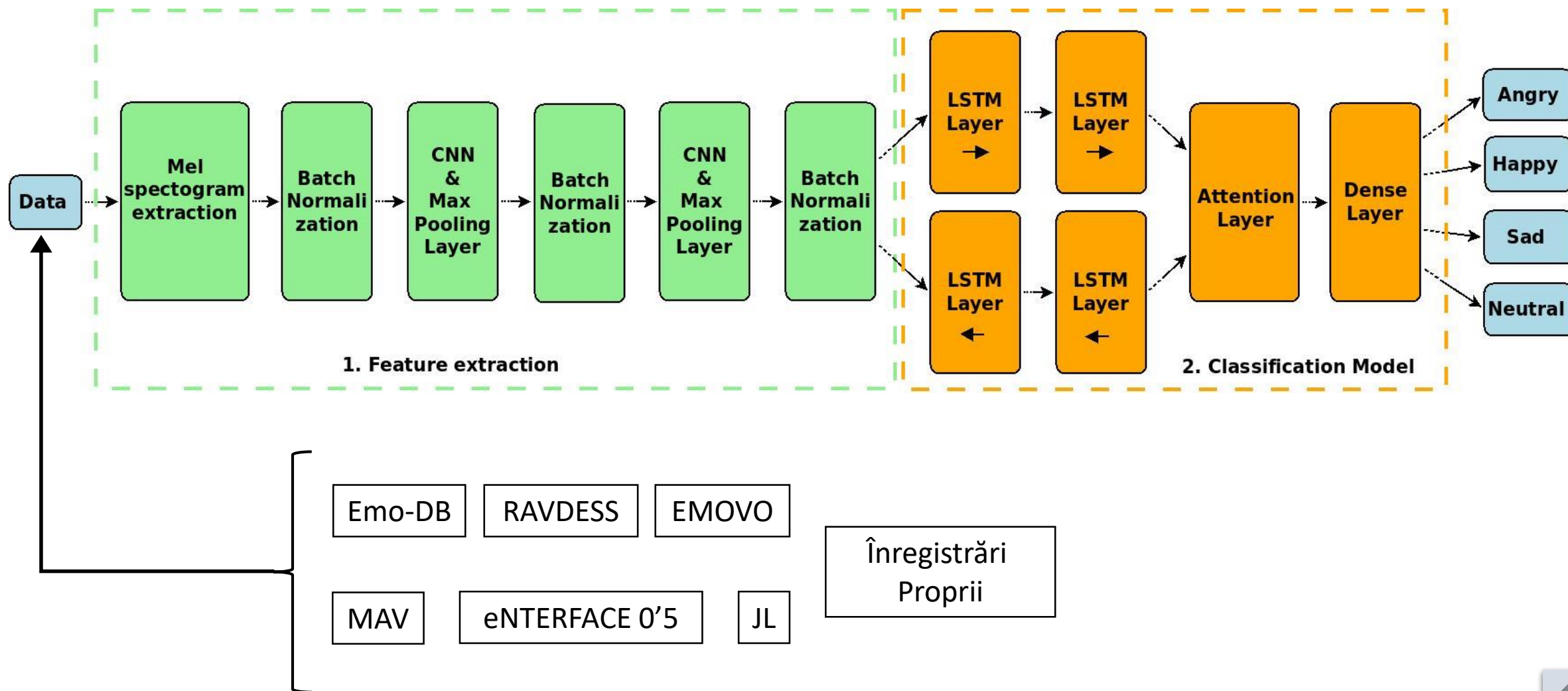
- Numărul redus de exemple de antrenare.
- Înregistrări jucate.
- Validarea .

Caracteristicile de intrare

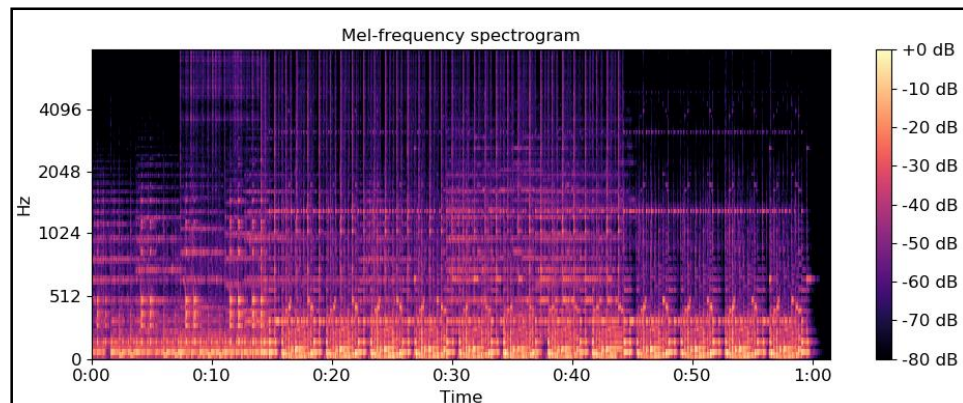
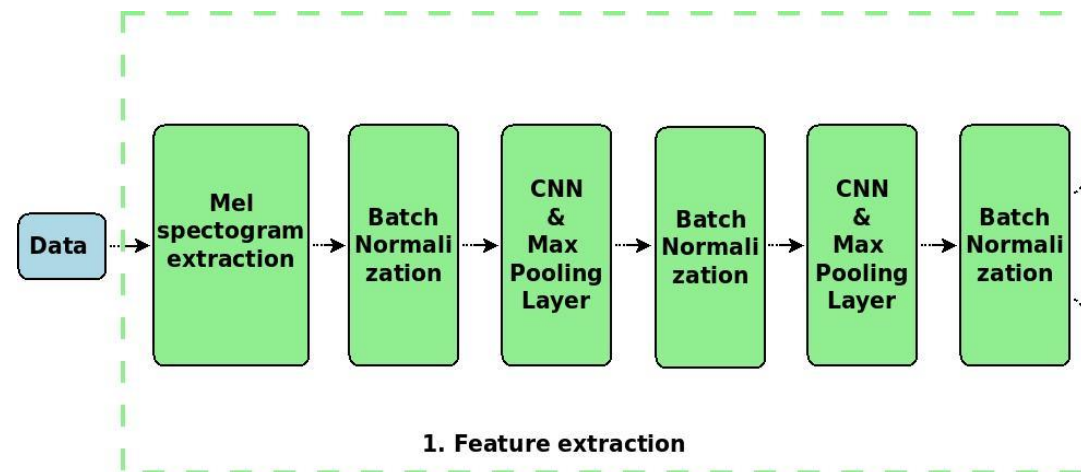
- Lipsa unui set de caracteristici de intrare reprezentative.
- “Hand-crafted”, necesită experți din domeniul audio.
- “End-to-end”=“black-box”.



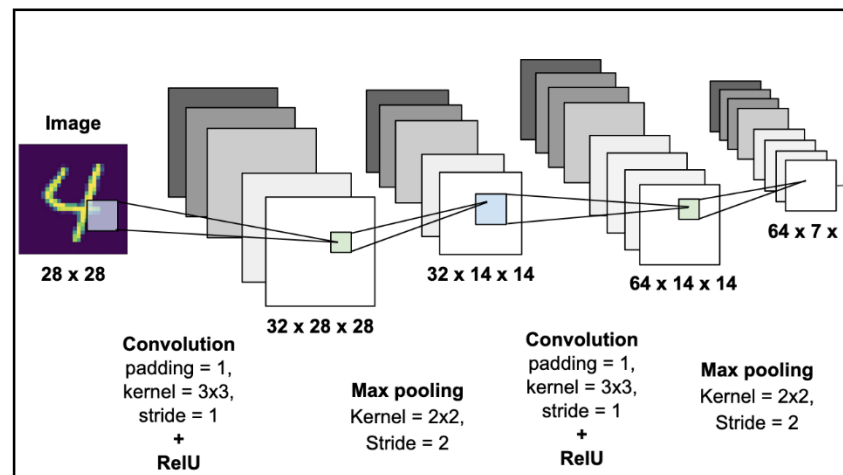
Soluția propusă



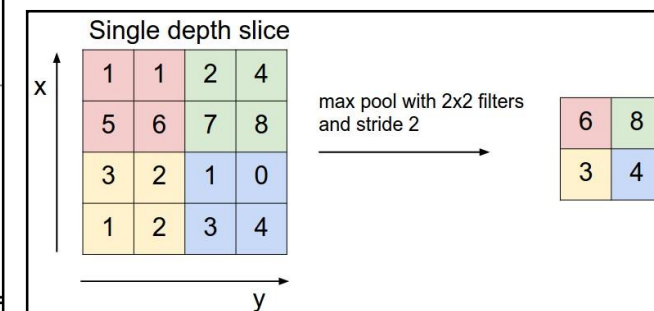
Extragerea caracteristicilor de intrare



Sursa [5]



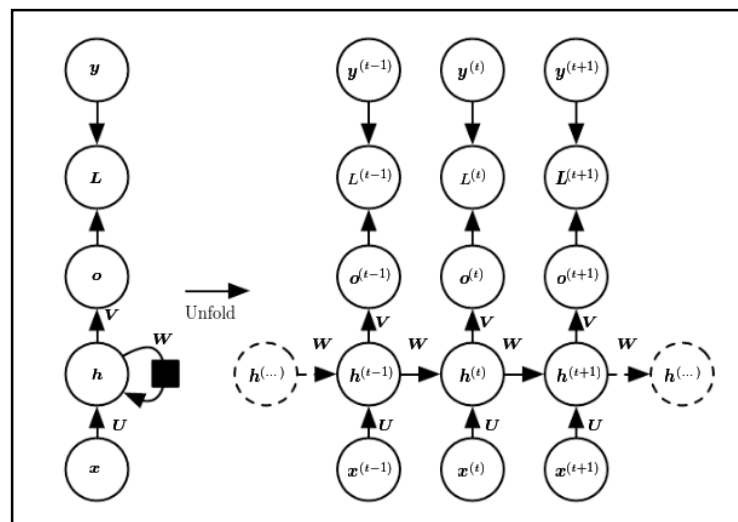
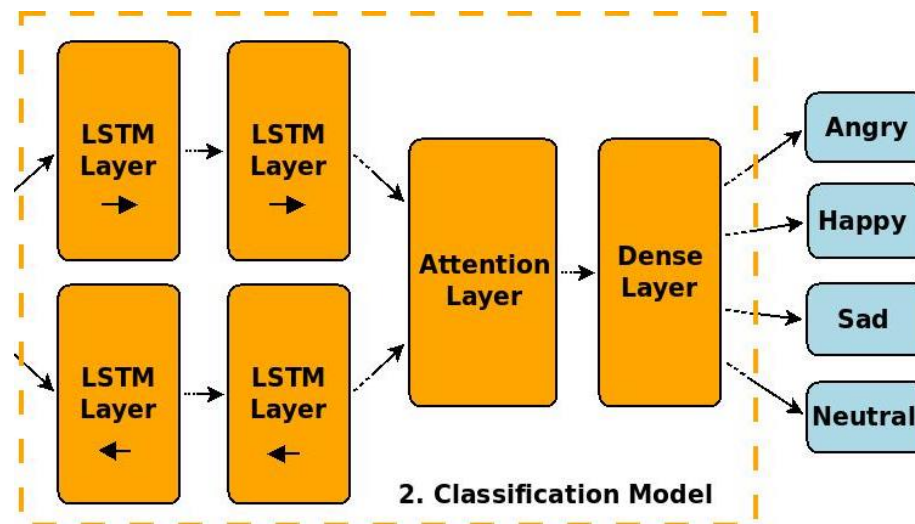
Sursa [6]



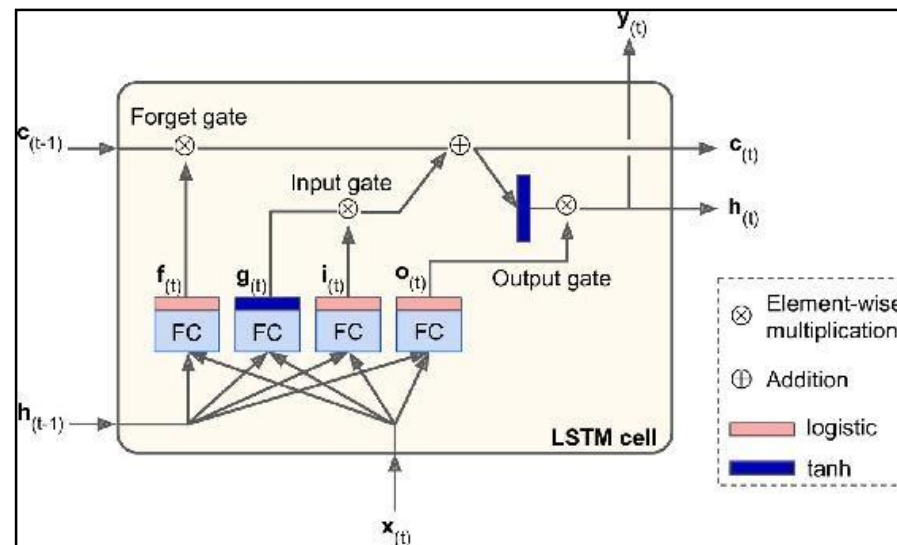
Sursa [7]



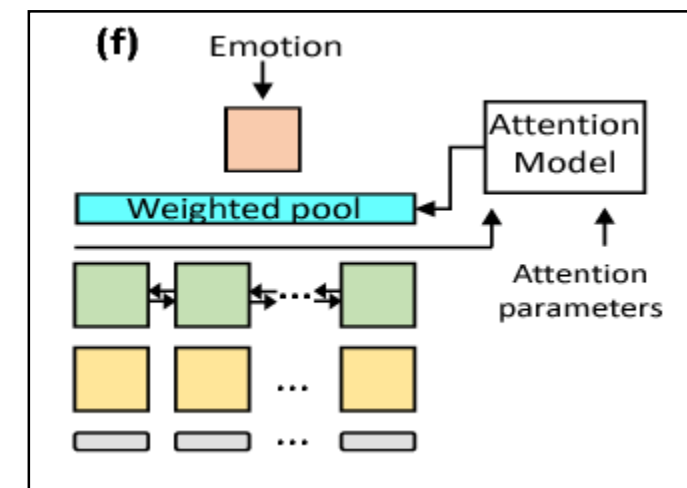
Modelul calificador



Sursa [8]



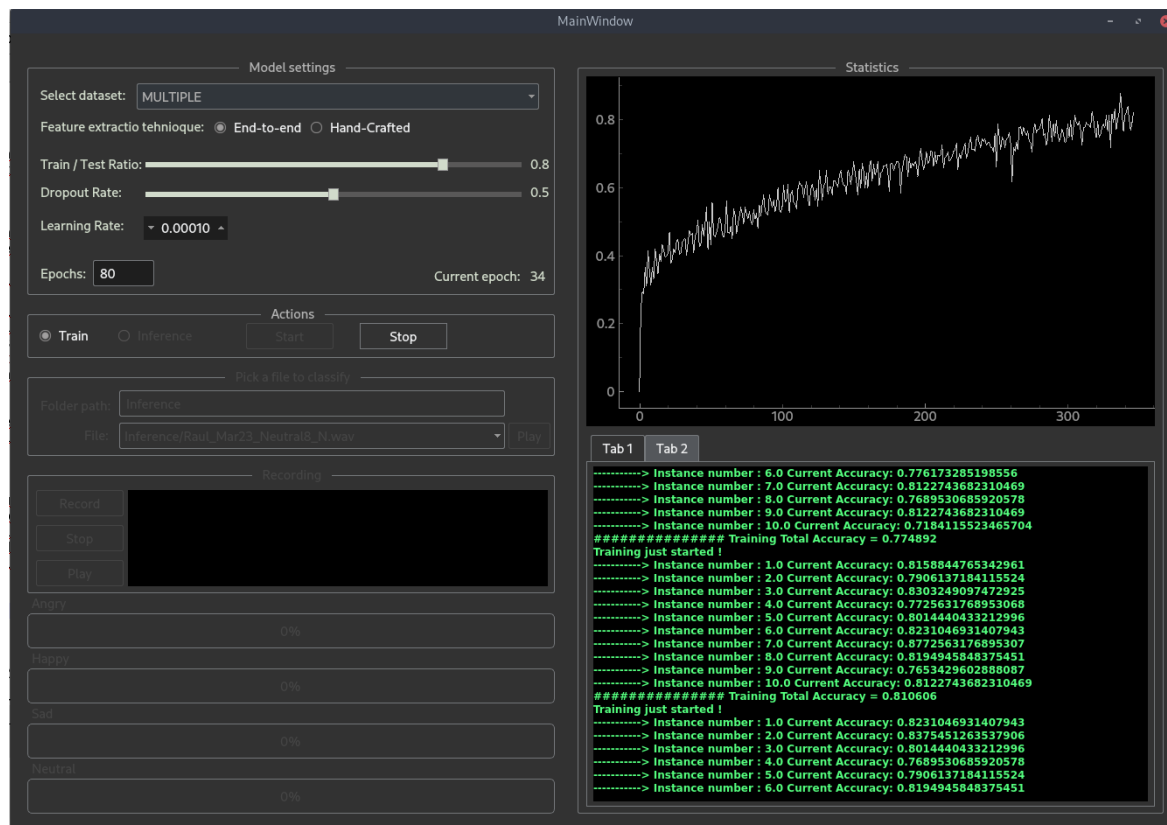
Sursa [9]



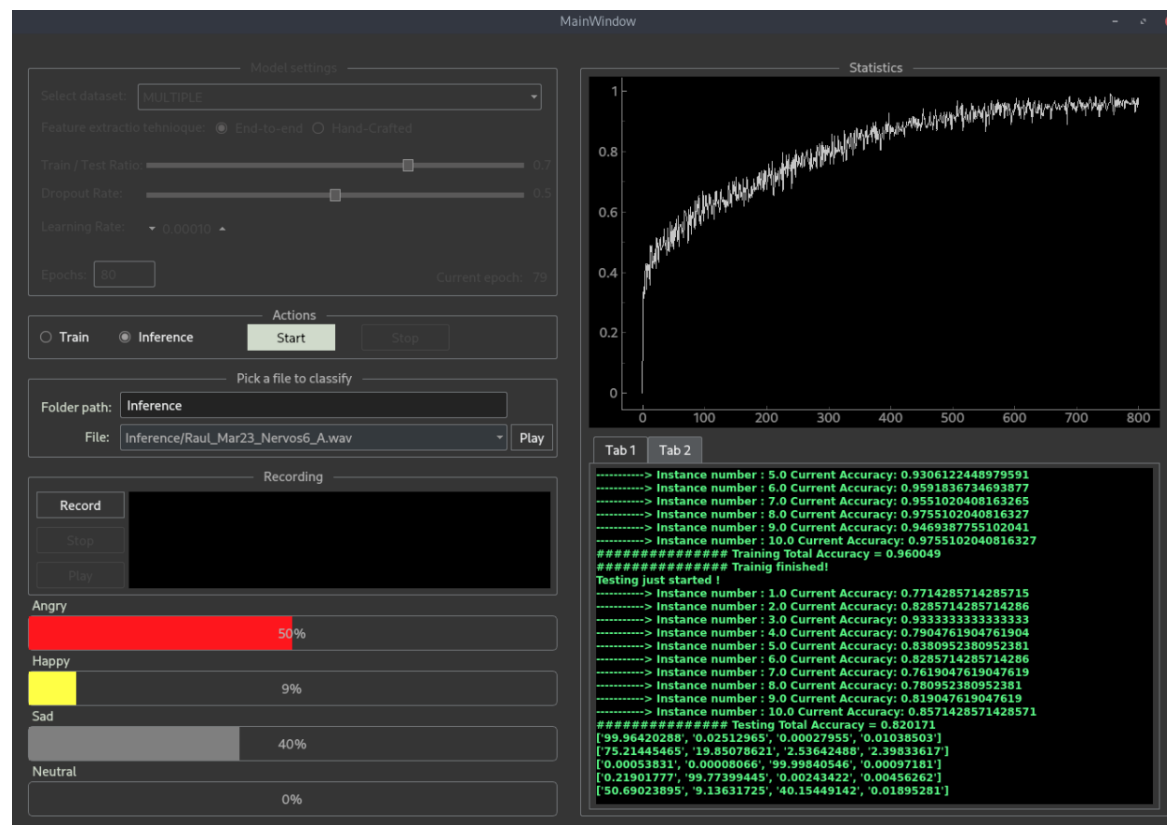
Sursa [4.1]



Interfața grafică



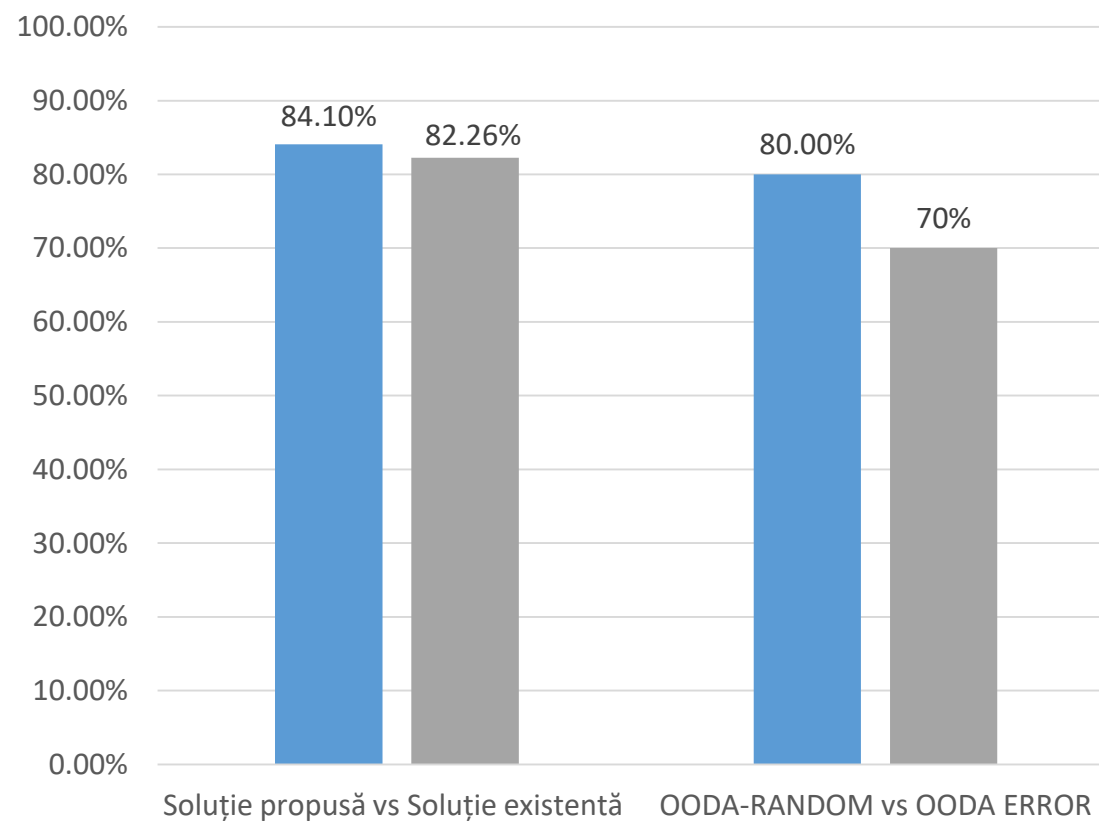
Interfața grafică în modul de antrenare



Interfața grafică în modul de inferență

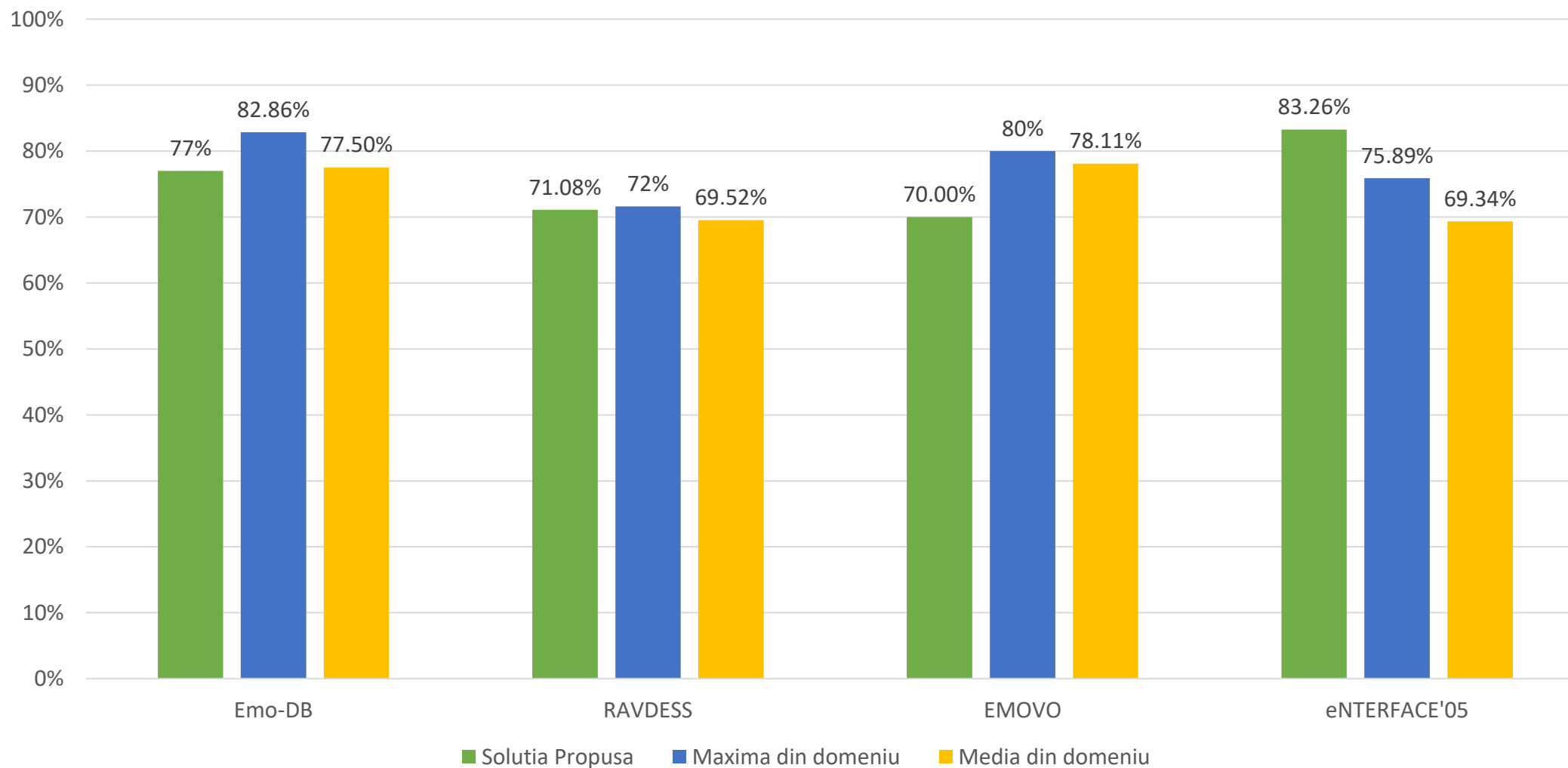


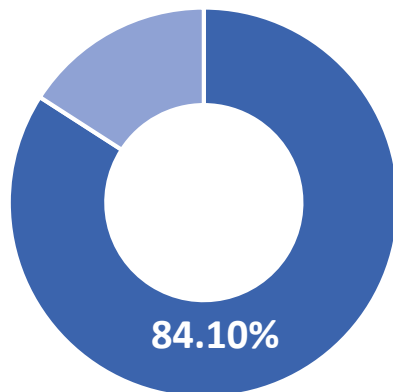
Rezultatele obținute pe intregul set de baze
de ate



Rezultate si experimente

Rezultatele individuale pe bazele de date





Concluzii

- Rezultate comparative cu cele de top din domeniu.
- Generalitatea și acuratețea obținută fac că soluția propusă să fie o fundație cu un potențial ridicat în combinație cu îmbunătățiri viitoare.

Posibile îmbunătățiri

- Mărirea numărului de baze de date.
- Diminuarea diferențelor dintre bazele de date.
- Antrenarea "multi-task".
- Combinarea soluției propuse cu un algoritm de detecție a emoției vizuale.

Vă mulțumesc pentru atenția acordată

- [1] https://www.brainbits.net/media/f904c10e0790ffd73cd6d64578852310/in-their-own-words-pt-1_featured%403x-1560x760.png
- [2] https://images.assetsdelivery.com/compings_v2/microone/microone1808/microone180800775.jpg
- [3] Frank Dellaert et al. “Recognizing Emotion In Speech”. In: *International Conference on Spoken Language Processing, ICSLP, Proceedings 3* (Dec. 1996).
- [4] Slide 3: 1. <https://www.researchgate.net/publication/336935196/figure/fig2/AS:820240816537601@1572572058375/Hard-parameter-sharing-for-Multi-Task-Learning-integrated-in-neural-networks.ppm>
 - 2. S. Mirsamadi et al. “Automatic speech emotion recognition using recurrent neural networks with local attention”. In: *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2017, pp. 2227–2231.
 - 3. http://practicalcryptography.com/media/miscellaneous/files/vadimg_1.png
 - 4. https://www.pyimagesearch.com/wp-content/uploads/2020/02/keras_autoencoders_applications.png
 - 5. https://www.researchgate.net/figure/Artificial-neural-network-architecture-ANN-i-h-1-h-2-h-n-o_fig1_321259051
- [5] https://librosa.github.io/librosa/_images/librosa-feature-melspectrogram-1.png
- [6] https://miro.medium.com/max/3744/1*SGPGG7oeSvVIV5sOSQ2iZw.png
- [7] <https://cs231n.github.io/assets/cnn/maxpool.jpeg>
- [8] Ian Goodfellow et al. *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press, 2016.
- [9] Aurlien Gron. *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. 1st. OReilly Media, Inc., 2017. isbn: 1491962291.