

Detectia tipului de motor prin analiza sunetului

Gherman Sebastian-Costin
343C2

I. ABSTRACT

In aceasta lucrare am experimentat diferite metode prin care se poate determina tipul carburantului folosit intr-un motor cu combustie interna prin analiza sunetului produs, atât la ralanti, cat si in sarcina.

Algoritmul este gândit sa facă o selecție binara intre motoarele Diesel (motorina) si motoarele pe benzina, atat aspirate natural, cat si supra-alimentate (turbo-suflanta).

II. INTRODUCERE

In zilele moderne, automobilele sunt un mijloc de transport foarte folosit, atat pentru transportul de oameni, cat si pentru transportul de marfa sau alte utilitati.

Ideea din spatele articolului pe care este bazat acest proiect (Spectral features for audio based vehicle and engine classification) este utilizarea sunetului pentru a captura date despre trafic. Spre deosebire de monitorizarea video, monitorizarea audio are mai multe beneficii.

1. Consum redus de resurse, atat din punct de vedere al stocarii datelor, cat si financiar (pretul unui microfon, fata de o camera infrarosu)
2. Un microfon poate inregistra si noaptea sau pe timp nefavorabil, fara a fi afectat in vreun fel, spre deosebire de o camera, care are nevoie de un senzor foarte scump pentru a avea o detectie buna pe timp de noapte.
3. Microfoanele pot sa fie instalate mult mai usor spre deosebire de o camera si pot sa fie pozitionate in asa fel incat sa nu distraga soferii
4. Un avantaj pe care sunetul il aduce este faptul ca se pot determina tipurile de motoare cu combustie interna folosite, acest aspect fiind imposibil utilizand doar canalul video.

Totusi, exista si dezavantaje in utilizarea canalului audio pentru a determina traficul si intensitatea lui. In situatii in care 2 motoare similare ajung sa fie suprapuse, este greu spre imposibil in a determina ca sunt 2 surse diferite. De asemenea, in situatii de trafic ingreunat, un singur microfon nu va putea detecta tot traficul, asa ca ar fi nevoie de mai multe, amplasate strategic, ceea ce ar ridica mult costul si ar elimina toate avantajele pe care le aduce in fata camerei video. Un alt aspect important este faptul ca o masina electrica, pana la o viteza aproximativ de 30 km/h

nu produce un sunet suficient de puternic pentru a putea fi detectata.

III. DATASET

Datele colectate in acest proiect au fost preluate din multiple surse. Clasificarea a fost facuta dupa cele 2 tipuri de motoare (diesel, benzina).

Sursele folosite pentru a captura datele au fost:

1. Youtube
2. Euro Truck Simulator 2
3. Sunete preluate manual

Sunetele preluate digital au fost exportate in format .wav pentru a pastra un nivel redus de compresie, si automat un nivel ridicat de fidelitate.

Pentru categoria diesel, au fost capturate aproximativ 7 minute de date, iar pentru categoria benzina 6 minute, in multiple segmente audio. Aceste segmente audio de 5-40 secunde au fost ulterior impartite in segmente de cate o secunda si au primit un label pentru categoria din care fac parte.

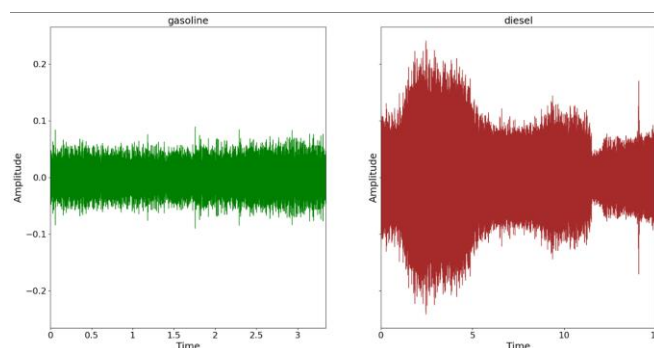


Fig. 1. In aceasta figura se pot observa cele 2 plot-uri pentru 2 sunete din categorii diferite. Prin simpla comparatie a celor 2 sunete, nu se poate determina cu acuratete tipul de motor, mai ales prin prisma faptului ca sunetele nu au fost inregistrate intr-un mediu controlat. Singura diferenta majora care poate sa fie detectata ar fi intensitatea sunetului motorului Diesel care este mult mai mare.

Pentru a determina cu acuratete diferentele dintre semnale, este necesara decompunerea sunetelor in sume de sinusoidale. Pentru acest lucru am folosit transformata Fourier. Folosind

FFT am putut determina in urmatorul grafic intensitatea sunetului pentru spectrul de frecvente.

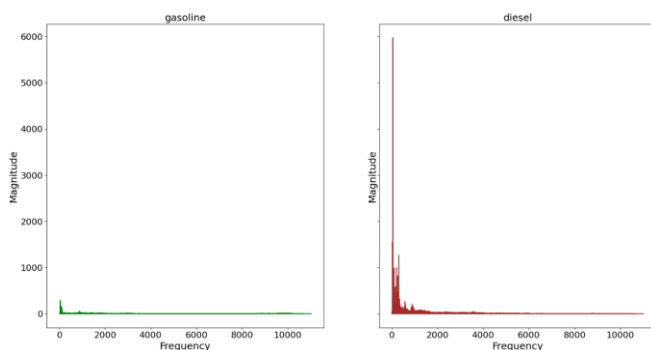


Fig. 2. Transformata Fourier aplicata pe cele 2 sunete plotata in functie de frecventa si magnitudine.

Un dezavantaj al utilizarii transformatei Fourier este faptul ca este pierdut domeniul timp. Avand in vedere ca timpul este un factor critic in determinarea tipului de sunet, suntem nevoiti sa folosim o alta transformata si anume Short Time Fourier Transform, pe scurt STFT. Avantajul acestei transformate este ca pastreaza domeniul timp, deoarece FFT-ul este aplicat pe o fereastră de dimensiune fixa care este miscata la dreapta cu un numar fix. Rezultatul aplicarii STFT se numeste o spectrograma.

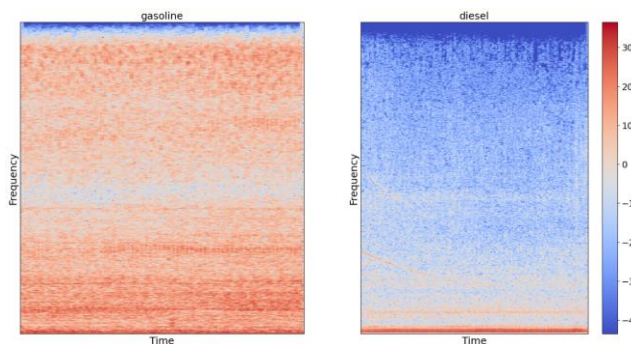


Fig. 3. Spectrograma realizata in urma aplicarii STFT pe cele 2 sunete. Fereastră are o dimensiune $n_{fft} = 2048$ iar shiftarea la dreapta se face cu un $hop_length = 512$. Se poate observa faptul ca in plot-ul pentru benzina, exista foarte mult zgomot, spre deosebire de cel pentru motorina. Astfel putem observa ca eliminand o parte din zgomot, motorul pe motorina este mai amplificat in frecventele joase.

Totusi, pentru a extrage „timbrul” motorului, o spectrograma nu este suficienta pentru ca nu ofera suficienta informatie algoritmului de Deep Learning. Pentru a combate acest lucru, folosim un concept numit **Mel Frequency Cepstral Coefficients (MFCCs)**. Timbrul sau textura unui sunet nu sta doar intr-o functie de frecventa, amplitudine sau tonalitate. Astfel, MFCC poate sa aproximeze auzul uman.

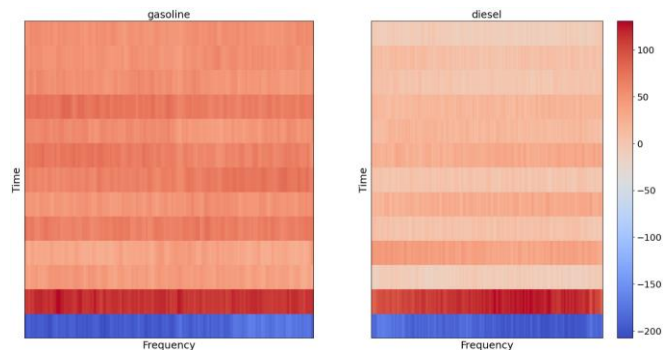


Fig. 4. Plotul obtinut in urma aplicarii algoritmului MFCC.

In continuare, pentru a forma datasetul pentru algoritmul de Deep Learning, este nevoie sa formam cate un grafic MFCC pentru fiecare secventa de 1 secunda al fiecarui sunet raw.

IV. RETEAUA NEURALA

Fiecare grafic MFCC este de fapt o matrice, ceea ce nu ajuta daca dorim sa introducem aceste date. Pentru a rezolva problema trebuie sa folosim un vector, astfel trecem matricea in format 1-D.

Primul layer este cel de flattening pe care il vom configura cu 572 neuroni, deoarece avem 13 coeficienti pe o coloana si 44 coloane care reprezinta transformatele Fourier. Avand in vedere ca avem un frame size de 2048 si un hop length de 512 si un sample rate de 22050 sample pe secunda, determinam ca in total sunt 44 de transformate Fourier.

Ulterior definim cele 3 layere ascunse cu 512, 256 si 64 de neuroni. Pentru a evita overfitting, folosim un dropout de 0.3 sau 30%. Avand un output format din doar 2 clase, vom folosi un loss function numit `binary_crossentropy`.

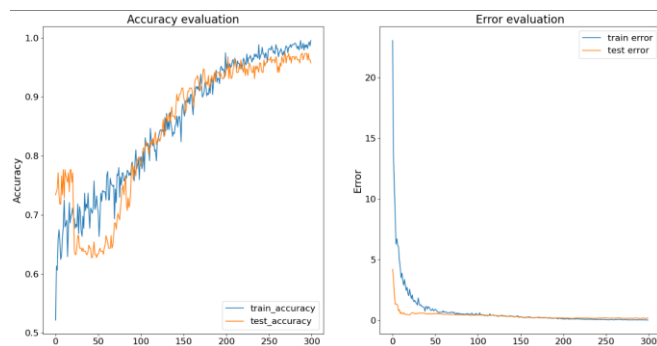


Fig. 5. Rezultatul evaluarii algoritmului in urma a 300 epochs.

De asemenea, algoritmul a fost testat si cu sunete din ambele clase care nu au fost introduse in datasetul de training. Rezultatele sunt mai mari decat asteptate, avand in vedere dimensiunea redusa a datasetului.

```

Running: BUILDING DATASET PHASE
Finished: BUILDING DATASET PHASE
Running: TESTING PHASE
Testing predictions
Diesel Successfully Prediction Rate: 100.00%
Gasoline Successfully Prediction Rate: 95.56%
Finished: TESTING PHASE

```

Fig. 6. Cel mai bun rezultat al ratei de predictie realizat pe 3 teste pentru fiecare categorie.

Engine Type / Success Rate	Diesel	Gasoline
Mean Accuracy	83%	60%
Min Accuracy	70	40
Max Accuracy	95	80
Mean Recall	93%	78%
Min Recall	85	60
Max Recall	100	95

Fig. 7. Rezultatul algoritmului in urma testarii.

V.RELATED WORK

Detectia tipului de motoare si a traficului exista de ceva timp, iar mai nou, cu ajutorul metodelor avansate de clasificare, folosind inteligenta artificiala se pot atinge nivele impresionante de acuratete.

Diverse clasificatoare au fost folosite de-a lungul timpului, cum ar fi SVM (support vector machines) pentru a determina tipul automobilului (automobil mic, autoutilitara, camion). Rata de succes a SVM este de aproximativ 80%.

VI. CONCLUZIE

Utilitatea reala a acestui mecanism de clasificare si detectie poate fi folosita in multe situatii chiar ca un ajutor pentru camerele de supraveghere. O varianta pentru a reduce din consumul de resurse la o camera conventionala de monitorizare ar fi sa se utilizeze impreuna cu un telefon pentru a determina daca este nevoie sa inregistreze la momentul respectiv. Daca s-ar folosi un alt mecanism de detectie al miscarii, ar fi destul de complicat sa detecteze miscarea fals pozitiva (cum ar fi animale sau oameni).

Un alt caz de utilitate ar fi in zonele izolate unde conectivitatea este limitata (curent / internet), astfel intervenind solutia microfonului. Pentru a determina, de exemplu, numarul de autocamioane care au trecut intrarea/iesirea dintr-un santier izolat, se poate folosi un microfon, iar rezultatele ar putea sa fie transmise atat prin tehnologie LTE cat si prin Wi-Fi Long Range, avand in vedere cantitatea redusa de informatie necesara pentru a realiza o clasificare si detectie a unui vehicul. De asemenea, este redusa foarte mult energia necesara pentru a utiliza un astfel de dispozitiv.

Asadar, situatiile prezentate reprezinta doar cateva dintre ariile care ar putea fi acoperite doar de un microfon sau de un set microfon/camera si un algoritm de detectie cu posibilitate de detectie a sunetului de motor, a tipului de motor sau a tipului de vehicul.

REFERENCES

- [1] Spectral features for audio based vehicle and engine classification, Alicja Wieczorkowska, Elzbieta Kubera, Tomasz Słowik & Krzysztof Skrzypiec, Journal of Intelligent Information Systems volume 50, <https://link.springer.com/article/10.1007/s10844-017-0459-2#Sec11>
- [2] Signal Processing — Engine Sound Detection, Paul Mora, <https://becominghuman.ai/signal-processing-engine-sound-detection-a88a8fa48344>
- [3] Deep Learning (for Audio) with Python, Valerio Velardo, <https://www.youtube.com/playlist?list=PL-wATfeyAMNrtbkCNsLcpoAyBBRJVZVInf>
- [4] Simple audio recognition: Recognizing keywords, https://www.tensorflow.org/tutorials/audio/simple_audio
- [5] <http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/>