

# Guitar HandTracking

\*Recunoaștere note muzicale

Butnaru Raimond-Eduard

Universitatea Tehnică "Gheorghe Asachi" din Iași

Iași, România

raimond-eduard.butnaru@student.tuiasi.ro

Guriuc Vlad Ionuț

Universitatea Tehnică "Gheorghe Asachi" din Iași

Iași, România

vlad-ionut.guriuc@student.tuiasi.ro

**Abstract**—Chitara este unul dintre cele mai populare și cele mai folosite instrumente. Popularitatea chitării o face unul din cele mai evidente alegeri pentru toți cei care își doresc să învețe să cânte la un instrument. Având în vedere popularitatea instrumentului, evident, putem discuta și despre gama largă de materiale online ce ajută la învățare și la perfecționare. Însă un mare dezavantaj al materialelor aflate în mediul online este lipsa flexibilității. Pe lângă aceasta, devine tot mai greu procesul de învățare online fără ajutorul unor profesioniști, de care nu toți dispunem. În acest document propunem o metodă de recunoaștere a notelor, respectiv a acordurilor cântate la chitară acustică/clasică. Metoda prin care ne-am propus să realizăm proiectul, exclude utilizarea procedurilor de machine learning pentru a crește portabilitatea produsului final.

## I. INTRODUCERE

Procesul de învățare a unui nou instrument, cum este și chitara, poate fi interesant și plăcut pentru pasionații de muzică, din nefericire acest proces devine greoi și neplăcut în lipsa ajutorului unui chitarist profesionist. Chitaristul este cel responsabil pentru felul în care începătorii învață acordurile, chitaristul, ar trebui să corecteze greșelile pe care le face începătorul și să îi analizeze performanțele. Însă lecțiile de chitară pot fi destul de costisitoare și pot consuma destul de mult timp. Alternativa, ar fi cursuri de chitară online sau aplicații mobile care facilitează procesul de învățare, dar până și acestea au dezavantajele lor. Sistemul de cursuri online pentru chitară nu este metoda ideală de a învăța deoarece nimeni nu poate corecta greșelile celor ce participă în acest curs, ceea ce face ca feedback-ul interactiv să nu existe deloc. Posibilitatea de a învăța prin aplicații mobile vine și cu un set de dezavantaje, unul dintre cele mai mari dezavantaje sunt lipsa instrumentului propriu-zis, unele aplicații forțează utilizatorul spre a folosi instrumentele virtuale care sunt doar simulări ale variantelor reale, altele se folosesc de microfonul telefonului pentru a testa acuratețea, dar în funcție de alte zgomote din jur acuratețea aplicației poate scădea drastic, iar restul aplicațiilor permit doar un număr de lecții pe zi, devenind o povară și făcând în moduri tot mai insistente reclame la un abonament care ar scăpa de reclame și de limitele de lecții pe zi. Scopul acestei lucrări este a facilita un nou proces de învățare mult mai rapid și mai eficient care aduce și feedback interactiv.

Prin proiectul nostru, ne propunem o alternativă de facilitare a procesului de învățare eficientă, folosind camera integrată în laptop-ul personal, aplicația va capta cadre ale feed-ului

live și va recunoaște acorduri și eventualele imperfecțiuni în execuția acestora. O posibilă extensie a acestei aplicații, ar fi încărcarea de videoclipuri cu recunoaștere în timp real, ajungând la generare automată de partituri.

## II. COMPONENTELE UNEI CHITARI

Fie chitara clasică, electrică sau acustică, compoziția acestora este similară indiferent dacă aceasta este un instrument "rece" sau unul ce necesită cabluri. Există totuși mici diferențe între aceste tipuri. Spre exemplu chitara clasică are un corp mult mai mic decât celelalte grupuri griful este mai lat, iar corzile, de cele mai multe ori sunt din nylon respectiv matase. Chitara acustică are bratul mai subțire dar este mai lungă și are corzile din metal. Chitara electrică este diferită doar datorită posibilităților de modificare a sunetelor prin ajutorul unui sistem de efecte.

Începând de sus, chitara cuprinde următoarele componente:

- Capul chitarei - acesta se află în varful chitarei
- Cheite de acordaj - fie ele din metal sau plastic, acestea pot fi strânse, respectiv slabite pentru a acorda chitara
- Pragusul - fie acesta fabricat din os, lemn, plastic, metal sau grafit are de obicei o culoare albă, ajută la înălțarea corzilor în partea superioară a chitării dar și la menținerea lor într-un loc stabil
- Tasta sau fret - sunt secțiunile de pe bratul chitarei
- Gatul, bratul sau griful - indiferent de termen acesta se referă la cea mai lungă secțiune a chitarei
- Butucul - acesta se află în spatele chitarei, este o bucată de lemn rezistentă care se prinde de partea inferioară a chitărei și de bratul acesteia pentru a nu permite corzilor să îndoie, respectiv să rupă griful de la tensiune.
- Rozeta - Are, de obicei, forma de cerc și se află pe mijlocul corpului chitărei, aceasta face legătura dintre corzi și cutia de rezonanță a chitărei
- Eclisa - forma unei chitari nu este intamplatoare, spațiul rotund de la mijlocul corpului chitărei este utilizat pentru poziționarea corectă a acesteia pe piciorul utilizatorului.
- Calus - calusul, ca și pragul, are scopul similar de înălțarea corzilor și de a le menține stabile doar ca acesta este mai gros, mai înalt și se află în partea inferioară a chitărei.
- Cordar - în figura 1 se poate observa un cordar bazat pe ace, acele puncte albe sunt ace groase din lemn care țin capatul corzilor stabil. La alte chitari, de regulă clasice,

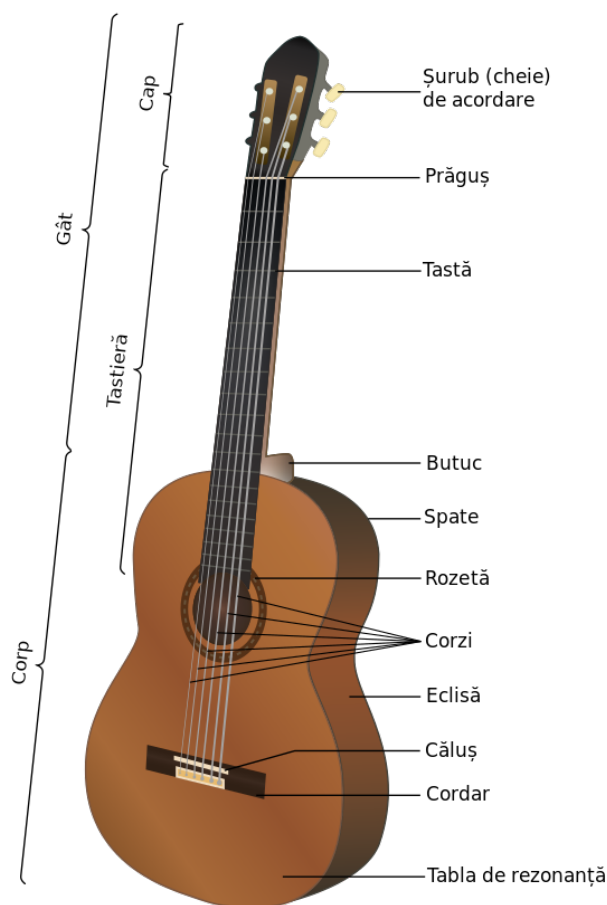


Fig. 1. Componentele chitarii

cordarul nu conține aceste ace iar corzile se leagă de cordar. La chitarile electrice cordarul se extinde în fața până în spate, iar corzile se inserează prin spatele chitarei.

- Tabla sau cutia de rezonanță - este cea mai mare parte a chitarei și practic în această zonă se produc sunetele pe care le produce instrumentul.

#### A. Motivație

Există grupuri foarte mari de oameni care își doresc să învețe să cânte la chitară. Tot acest grup de oameni încep să își piardă motivația când procesul de învățare devine unul foarte greu și neplăcut. Cel mai greu este pentru cei care nu merg la vreun curs de chitară sau au vreun ghid profesionist, pentru această categorie de persoane se pot crea întreprinderi greșite care rămân în timp neobservate.

#### B. Problema

Înainte ca cineva să înceapă vreun curs de specialitate pentru chitară, sigur a căutat cândva pe YouTube: *How to play the guitar* și au descoperit o sumedenie de clipuri care nu au nicio legătură între ele și au în comun același principiu: "Reproduce ceea ce vezi chiar dacă nu știi cât de corect este". De menționat este și posibilul zgomot des întâlnit în videoclipuri, care poate crea discomfort sau poate presa utilizatorul spre a repeta

aceeași secțiune a clipului până când aude ceea ce crede el că ar trebui să audă, proces care consumă foarte mult timp de altfel. Este nevoie de un sistem portabil care face învățarea acestui instrument accesibilă pentru toate lumea, oricând și oriunde.

#### C. Cum se formează acordurile

Acordurile se formează doar pe brațul chitării, urmând teoriile muzicale, avem 12 acorduri de bază, majore și minore, cu forme unice ce pot fi executate pe intervalul de taste 1-5, ele sunt cunoscute și ca "Base chords". Mai departe pe gâtul chitării avem "Barre chords", aceste acorduri au forma mâinilor foarte similare între ele, însă diferă poziția acestora pe gâtul chitării. Spre exemplu, acordul Fa major (F), de regulă se află între tastele 1 și 3, însă, executat pe intervalul de taste 3-6 acesta devine Sol major (G).

### III. CONCEPTE ASEMĂNĂTOARE

În lucrarea lui Yingdong Ru intitulată *Computer Assisted Chord Detection Using Deep Learning and YOLOV4 Neural Network Model*[1] este prezentată detectarea acordurilor de pe partituri folosind machine learning, predisponând aplicația la diferite seturi de date care aveau de recunoscut unde se află notele pe portativ. Un alt articol asemănător este lucrarea lui ALbert Mitjans Coma intitulată *Visual recognition of guitar chords using neural networks*[2]. În acest articol este pusă în vedere presupunerea mai multor acorduri în urma detectării mâinii, urmând mai apoi intersectarea cu rezultatul cel mai apropiat de adevăr, adică cel mai credibil. Lucrarea cea mai asemănătoare este *GuitarGuru: A Realtime Guitar Chords Detection System* [3]. În această lucrare, într-adevăr, se folosește machine learning, însă apare doar în procedura de comparare, nu și în cea de recunoaștere. Recunoașterea acordurilor este realizată pe baza calculului formelor mainilor și a punctelor de discontinuitate aflate pe fiecare tastă a fiecărei corzi. Din moment ce este foarte importantă mimica mâinilor pentru detectarea acordurilor, următoarea lucrare se referă la un sistem de recunoaștere a mimicii mâinilor care captează gesturile făcute de către mâinile utilizatorului pentru a controla un sistem computerizat[4].

Lucrarea *Finger-String Interaction Analysis in Guitar Playing With Optical Motion Capture* [7] prezintă o metodă de urmărire și prioritizare a mainii. Plasând mai întâi punctele unde se află degetele pe chitară înainte de a începe alte prelucrări ale imaginii. Procedeu conține și captarea dintr-o imagine a vibrațiilor corzilor pentru crearea de partituri în urma încărcării unui videoclip. Lucrarea *Finger-String Interaction Analysis in Guitar Playing With Optical Motion Capture by Alfonso Perez-Carillo 2019*. [8] prezintă o metodă de recunoaștere a gesturilor mainii utilizând o rețea neuronală convoluțională (CNN), adaptată din LeNet-5. Imaginile mainii sunt preprocesate pentru eliminarea fundalului, filtrarea culorii, blurare, și extragerea conturului. Un detector de mană estimează centrul și raza palmei, iar regiunea de contur este redimensionată și centrată înainte de a fi introdusă în acest CNN.

Acest sistem este evaluat și arată o precizie înaltă în recunoașterea gesturilor mâinii, facilitând monitorizarea gesturilor într-un context de recunoaștere a mișcărilor. Lucrarea *Abhishek B., Kanya Krishi, Meghana M. "Hand gesture recognition using machine learning algorithms" 2020.*[9] descrie un sistem de recunoaștere ce a fost dezvoltat pentru capturarea gesturilor efectuate de un utilizator și gestionarea unui sistem informatic bazat pe informațiile receptate. Acest sistem urmărește să dezvolte un sistem de recunoaștere a gesturilor pe cale vizuală ce prioritizează performanța, cu o rată de detecție cât mai ridicată, și care poate funcționa în timp real într-un sistem de interacțiune om-computer în condițiile în care utilizatorul nu se află într-un mediu ce ar putea îngreuna considerabil condițiile de funcționare a sistemului (utilizarea de accesorii redundante precum manusile sau utilizarea unui fundal încărcat de obiecte sau persoane). Lucrarea *Julien Osmalskyj, Jean Jacques Embrechts, Marc Van Droogenbroeck, Sébastien Piérard "Neural networks for musical chords recognition" 2012.*[10] reprezintă un studiu realizat în vederea recunoașterii de acorduri muzicale utilizând o rețea neuronală feed-forward. Cu ajutorul unui set de date ce conține o serie amplă de înregistrări (2000 de acorduri de chitară) ale acordurilor într-o cameră anecoică și într-o cameră în care există zgomot se mărește acuratețea algoritmului. Metoda propusă are rezultate semnificative și poate fi aplicată și altor instrumente: pian, vioară și acordeon. Marja de eroare pentru aceste 4 instrumente este de 1-13 procente spre deosebire de alte metode a căror marja de eroare variază între 8 și 32 procente. Această metodă își poate găsi utilitatea în aplicația noastră în etapele finale de dezvoltare, când interacțiunea utilizator-computer va fi mult mai accentuată, deoarece utilizatorul va putea acorda chitara sau va putea testa corectitudinea notelor sau a pieselor muzicale cântate de el însuși. Lucrarea *State-of-the-Art Model for Music Object Recognition with Deep Learning*[11] propune un model de recunoaștere a portativelor și partiturilor tipărite și folosește o rețea neuronală. Sistemul este capabil să preia partiturile și să furnizeze felul, durata și înălțimea notelor. Soluția urmărește să atingă o acuratețe cât mai înaltă și este relevantă în contextul nostru deoarece, prin extensie, pune accentul pe identificarea notelor sau acordurilor de pe chitară pe care s-au aplicat anumite efecte, precum tremolo, staccato sau legato.

#### IV. METODA UTILIZATĂ

Metoda propusă spre implementare utilizează un algoritm de procesare a imaginilor captate dintr-un feed live, oferit de către cameră, iar acest algoritm se poate împărți în mai multe etape:

- Ștergerea fundalului și a secțiunilor redundante
- Recunoașterea și secționarea corzilor și a tastelor
- Plasarea coordonatelor degetelor pe brațul chitarii
- Recunoașterea acordurilor/notelor

##### A. Ștergerea fundalului și a secțiunilor redundante

Pentru a evidenția chitara în interiorul imaginii, vom elimina fundalul și alte "informații" redundante din imaginea captată.



Fig. 2. Imagine cu fundal eliminat

Metoda finală va trebui să elimine aceste informații inutile fără modificarea secțiunii principale, va trebui să se adapteze la variații de lumină și să aibă o complexitate redusă pentru a nu apărea desincronizări în timpul rulării aplicației. Pentru scopul nostru, am testat algoritmi Running average Gaussian, Mixture of Gaussians (specific OpenCV) și Frame difference. Frame difference este eficient dar schimbul fundalului poate duce la erori majore, fiind nevoie de actualizare frecventă a fundalului în manieră manuală pentru a reduce erorile.

Mixture of Gaussian folosește un set de imagini pentru a construi un model de extragere, fiind foarte precis și cu posibilitatea de actualizare a modelului în timp de real, acesta în schimb scade destul de mult performanțele oferite de aplicație.

Running average Gaussian [5] este eficient, suportă schimb de fundal, dar pierde informații în condițiile în care pixelii nu suferă variații observabile.

Un alt algoritm care elimină elementele redundante este Gaussian blur. Din moment ce vom face procesare pe imagini de tip grayscale, această metodă ne avantajează cel mai mult și ca timp dar și ca performanțe.

În această lucrare ne propunem să folosim metodele menționate anterior în funcție de care este mai avantajoasă pentru fiecare pas de prelucrare a imaginilor captate.

##### B. Recunoașterea și secționarea corzilor și a tastelor

Informațiile redundante fiind eliminate, aplicația se poate concentra asupra obiectului în discuție, chitara, fiind necesară segmentarea tastelor de pe brațul chitarii.

Pentru a recunoaște corzile vom aplica pe o imagine grayscale un filtru Sobel pe axa Y. În urma aplicării acestui filtru, folosind transformata Hough, vom detecta liniile orizontale. Mai departe decupăm imaginea pe secțiuni verticale la diferite puncte după care se calculează spațiile dintre corzi, mai precis spațiile pe axa Y.

În final se formează linii care să separe fiecare coardă în parte după care se returnează imaginea cu liniile desenate.

Pentru detectarea tastelor, logica din spate este similară, se aplică un filtru Sobel pe axa X după care se aplică iar transformata Hough. Aici în schimb, se fac niste presupuneri și se calculează niste decupaje. Pe baza pozițiilor acestor decupaje, vom calcula niste rapoarte, care vor încerca să

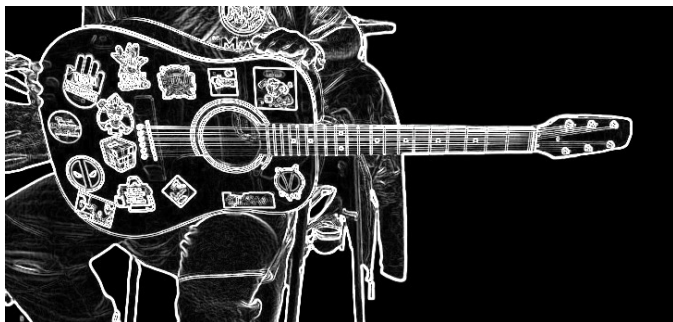


Fig. 3. Imagine cu filtru Canny aplicat



Fig. 4. Imagine cu filtru Canny decupată

deseneze pe imagine linii verticale ce vor arata pe unde s-ar afla tastele.

Pentru minimizarea erorilor ce pot apărea la filtrare, vom decupa brațul chitării din imaginea principală, concentrându-ne și mai mult pe chitară în sine. Cel mai mare obstacol pe care îl putem întâlni este iluminarea slabă, dacă lumina nu bate într-un mod specific pe chitară, există foarte mare șanse ca aplicația să nu vadă corzile sau să nu vadă tastierele.

În continuare, dorim să aplicăm o combinație între filtrul Canny și transformata Hough probabilistică. Pe baza liniilor orizontale create de transformata Hough, vom roti imaginea până când acestea ajung să fie perfect orizontale. Rotirea este realizată în urma aproximării unui centru pe brațul chitării, apoi sunt calculate punctele extreme care sunt rotite în jurul centrului până se ajunge la o poziție orizontală.

### C. Plasarea coordonatelor degetelor pe brațul chitarii

Modulul de HandTracking din biblioteca cvzone are integrat o metoda eficienta de detectare a degetelor, mai precis acest modul poate aproxima cu foarte mare acuratete si un set de puncte de la fiecare mana. Ba mai mult, modulul recunoaste



Fig. 5. Aplicarea transformatei Hough probabilistică

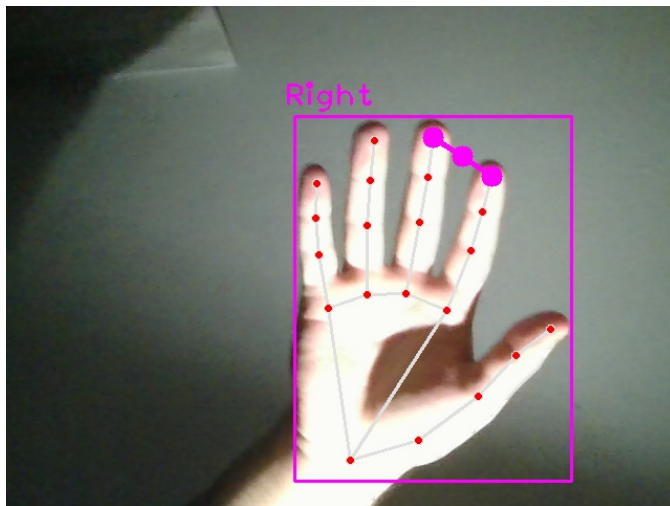


Fig. 6. Imagine cu mana captata

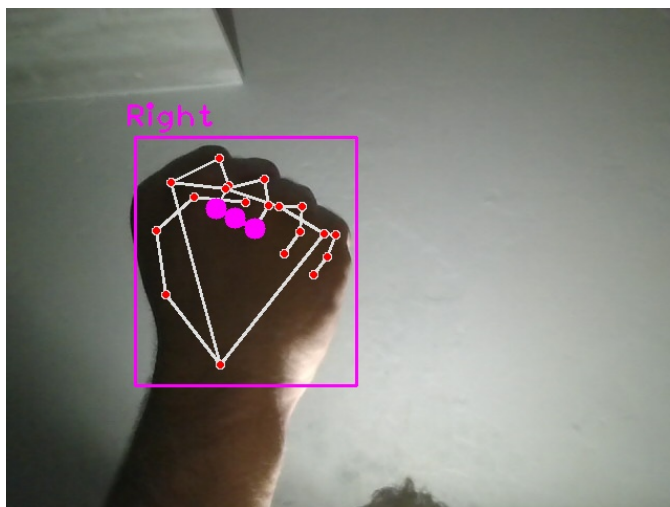


Fig. 7. Imagine in care degetele sunt obscurate

si ce mana se afla in cadru (mana stanga sau mana dreapta). Folosind o captura de frame intr-un punct oarecare al stream-ului live, se vor salva coordonatele varfurilor degetelor pe imaginea originala, dupa care aceasta va fi prelucrata pentru a cauta corzile si tastele.

**De mentionat:** modulul de HandTracking din biblioteca cvzone ofera si aproximari foarte precise care pot detecta pozitiile varfurilor degetelor chiar si daca acestea sunt obscurate de dosul palmei conform figurilor 6 si 7.

### D. Recunoașterea acordurilor/notelor

Pentru pasul de recunoaștere ar trebui să ajungem în prima faza cu imaginea pe care se afla trasate liniile corzilor, liniile tastelor și punctele desenate de la varfurile degetelor. Transmise în acest pas sunt și coordonatele și seturile de puncte care vor fi folosite pentru referințe. Într-un final, acordurile vor fi recunoscute verificând dacă coordonatele degetelor se



afla între patratele create dintre intersecția liniilor tastelor cu liniile corzilor.

## V. DESCRIEREA MODULELOR UTILIZATE

### A. *imageModule*

Scopul funcțiilor prezente în acest modul este unul simplu, dar vital. Aceste funcții au rolul de a returna un mesaj de eroare în cazul în care imaginea sau calea sa nu pot fi citite, de a returna imaginea dorită sau de a aplica imaginii date, după caz, un filtru Canny, Sobel, Laplace sau transformata Hough. Practic în acest modul descrie clasa imaginilor ce urmează a fi procesate.

### B. *Rotate and crop*

În acest modul regăsim o serie de funcții denumite sugestiv, ce ne ajută să rotim o imagine data astfel încât să putem aduce gatul chitării într-o poziție standard pentru a facilita identificarea acordurilor prin aplicarea ulterioară a filtrelor CannyEdge și a transformatei Hough. Odată ce imaginea este rotită, gatul chitării va fi decupat din imagine pentru a optimiza procesul de identificare. Pentru a asigura stabilitatea algoritmului am introdus o funcție de resize ce va redimensiona imaginile pe o rezoluție de maxim 1920x1080.

### C. *Watershed segmentation*

În prima instanță, acest modul convertește imaginea data ca parametru într-o imagine cu tonuri de gri ce va fi binarizată cu ajutorul binarizării Otsu/threshold. După ce erodăm (reducem numărul de obiecte din imagine) și dilatăm (extindem dimensiunea obiectelor rămase) imaginea, creăm o imagine de fundal. Combinăm imaginea de fundal cu cea dilatăta și vom obține marcatorii cu ajutorul cărora se va aplica filtrul Canny. După ce aplicăm algoritmul watershed pentru a segmenta imaginea, o binarizăm cu Otsu și apoi o inversăm. Aplicăm operația AND pe biți între imaginea originală, unde imaginea binară va fi mască și repetăm procesul dar de această dată mască va fi imaginea binară inversată. Bazându-ne pe aceste diferențe putem contura imaginea finală.

### D. *Grid detection*

Acest modul primește un parametru reprezentat de o imagine cu gatul chitării. În funcția "string detection" pe această imagine se detectează corzile folosind transformata Hough. Se creează o imagine cu corzile pe baza detecției și se delimitează imaginea în porțiuni verticale pentru a calcula spațiile dintre corzi. Calculează diferențe de pixeli pentru a identifica pozițiile unde pot fi prezente corzi și trasează linii orizontale în dreptul corzilor. Funcția "fret detection" primește ca parametru aceeași imagine și este folosită pentru a detecta tastele prin decuparea imaginii în segmente orizontale și calcularea spațiilor dintre ele. Odată ce aceste operații sunt terminate, se vor afișa linii verticale deasupra tastelor detectate.

## VI. DIFICULTĂȚI ÎNTÂMPINATE

De-a lungul realizării acestei aplicații am fost întâmpinați de probleme legate de implementarea sau compatibilitatea unor module, sau de probleme legate de utilizarea echipamentelor puse la dispoziție. Cu timpul, am reușit să trecem peste o parte din aceste impedimente, total sau parțial. Primele dificultăți au început să apară după rezultatele intermediare. Pentru început, detecția de corzi are șanse foarte mari de eșec la ultima, respectiv cea mai subțire coardă, recunoscând astfel 5 din 6 corzi, o soluție ar putea fi presupunerea în mod implicit a 6 corzi și căutarea continuă a exact 6 corzi. Mai apoi avem detecția tastelor de pe brațul chitării care are secțiuni desincronizate, soluție ar fi segmentarea pe spații albe. Ultima dificultate legată de implementarea modulelor, este detectarea vârfulor degetelor. Pe parcursul dezvoltării aplicației am constatat că este optim pentru noi să utilizăm camera integrată în laptop, întrucât camera ZED mini încetinește procesele de calcul, detecție și recunoaștere, ale programului realizat în python, aplicate pe imaginile capturate de aceasta.

## VII. BIBLIOTECA CVZONE

Asa cum am explicat anterior, unul din pașii necesari pentru recunoașterea acordurilor este recunoașterea poziției degetelor mâinii stângi pe tastierele de pe gatul chitării. Biblioteca cvzone are implementate diverse module ce realizează detecția și recunoașterea feței și mâinilor captate în diferite poziții din diferite unghiuri. Pentru aplicația noastră am utilizat modulul HandTracking. Acesta, cu ajutorul bibliotecii mediapipe este capabil să recunoască ambele mâini și exportă reperele drept pixeli. Ne folosim și de o serie de funcționalități puse la dispoziție de această bibliotecă. Funcția fingersUp preia printr-un parametru imaginea ce conține mâna, detectează dacă este stângă sau dreaptă, și detectează câte degete sunt ridicate. Funcția findDistance preia o imagine și două coordonate din aceasta și calculează distanța dintre ele, utilă în contextul nostru pentru determinarea distanței dintre taste și poziționarea degetelor pe tastieră. De asemenea, detecția realizată de modul poate fi marcată de încadrarea mâinii într-un patruleter.

## VIII. IMPARTIREA TASK-URILOR ECHIPEI

## IX. CONCLUZIE

Chitara este un instrument foarte popular și multă lume dorește să învețe să cânte la acesta. Dar evident, există dificultăți în procesul de învățare și există foarte puține platforme user-friendly care să faciliteze acest proces. De aceea am luat inițiativa de a dezvolta un sistem ce ajută chitariștii începători. În acest document am dezvoltat un sistem care urmărește modelul mâinii după care am capturat poziția varfurilor degetelor. Mai departe am preluat imaginea respectivă și am aplicat filtre Sobel și Canny pe imaginea transformată în grayscale pentru a putea aplica transformata Hough și pentru a calcula drepte ce se află pe lungimile corzilor, respectiv pe pozițiile tastelor.

TABLE I

ID Task	Descriere task	Membru echipa
1	Documentarea posibilitatilor camerei ZED Mini	Guriuc Vlad
2	Research pe filtrul Sobel	Butnaru Raimond
3	Documentarea aplicarii transformatei Hough	Guriuc Vlad
4	Documentarea diferitelor functii si librării opencv	Guriuc Vlad
5	Implementarea primelor variante de prelucrari si primele rezultate	Butnaru Raimond
6	Implementarea sectionarii si a decuparii imaginii	Butnaru Raimond
7	Filtrarea fundalului	Guriuc Vlad
8	Sedinta de discutie a rezultatelor intermediare	Guriuc Vlad si Butnaru Raimond
9	Implementarea detectiei corzilor si a tastelor	Butnaru Raimond
10	Documentare pe cv-zone si HandTracking module	Guriuc Vlad
11	Implementarea modului de HandTracking si portarea aplicatiei spre realtime	Butnaru Raimond
12	Update la documentatia si raportarea finala	Guriuc Vlad
13	Recoltarea de rezultate pentru prezentarea finala	Butnaru Raimond

## X. OBIECTIVE PE VIITOR

Ne propunem ca în viitor să ne reîntoarcem la această aplicație și să îi sporim portabilitatea și accesibilitatea, dar și să îi mărim repertoriul de acorduri la acorduri și mai complicate cum sunt cele cu septimă sau cele suspendate. În punctul curent operațiile sunt destul de lente și necesită o putere de procesare relativ mare pentru a putea duce la îndeplinire obiectivul.

## XI. RECUNOȘȚINȚĂ

Suntem recunoscători de această oportunitate și mulțumim Facultății de Automatică și Calculatoare a Universității Tehnice "Gheorghe Asachi" din Iași pentru suportul și pentru facilitățile oferite în cadrul dezvoltării proiectului.

## REFERENCES

- [1] Computer Assisted Chord Detection Using Deep Learning and YOLOV4 Neural Network Model by Yingdong Ru, 2021.
- [2] Visual recognition of guitar chords using neural networks by Albert Mitjans Coma and Artur Carnicer.
- [3] GuitarGuru: A Realtime Guitar Chords Detection System by Varsha Nagpurkar, Neha Pattankar, Tripti Nayak Allan D'Souza and Nipun Henriques, 2023.

- [4] B., Abhishek, Krishi, Kanya M., Meghana, Daaniyaal, Mohammed S., Anupama. "Hand gesture recognition using machine learning algorithm" Computer Science and Information Technologies 2020.
- [5] Cerman Martin : SUBTRACTION USING RUNNING GAUSSIAN AVERAGE : A COLOR CHANNEL COMPARASION
- [6] Lukas Tuggener, Thilo Stadelmann, Ismail Elezi, Jürgen Schmidhuber "Deep Watershed Detector For Music Object Recognition" 2018.
- [7] Finger-String Interaction Analysis in Guitar Playing With Optical Motion Capture by Alfonso Perez-Carillo 2019.
- [8] Pei Xu "A Real-time Hand Gesture Recognition and Human-Computer Interaction System" 2017
- [9] Abhishek B., Kanya Krishi, Meghana M. "Hand gesture recognition using machine learning algorithms" 2020.
- [10] Julien Osmalskyj, Jean Jacques Embrechts, Marc Van Droogenbroeck, Sébastien Piérard "Neural networks for musical chords recognition" 2012.
- [11] Zhiqing Huang, Xiang Jia and Yifan Guo "State-of-the-Art Model for Music Object Recognition with Deep Learning" 2019.