

ROZPOZNÁVANIE FAREBNÝCH NOTOVÝCH ZÁPISOV

ADAM DETKO

ŠKOLITEĽ: RNDR. ZUZANA BERGER HALADOVÁ, PHD.

PROBLÉM

- Zápis formou farebných nôt pre xylofón
- Rôzna farebnosť xylofónov

Ode To Joy (C Major)
Glockenspiel

L. van Beethoven

singing-bell.com

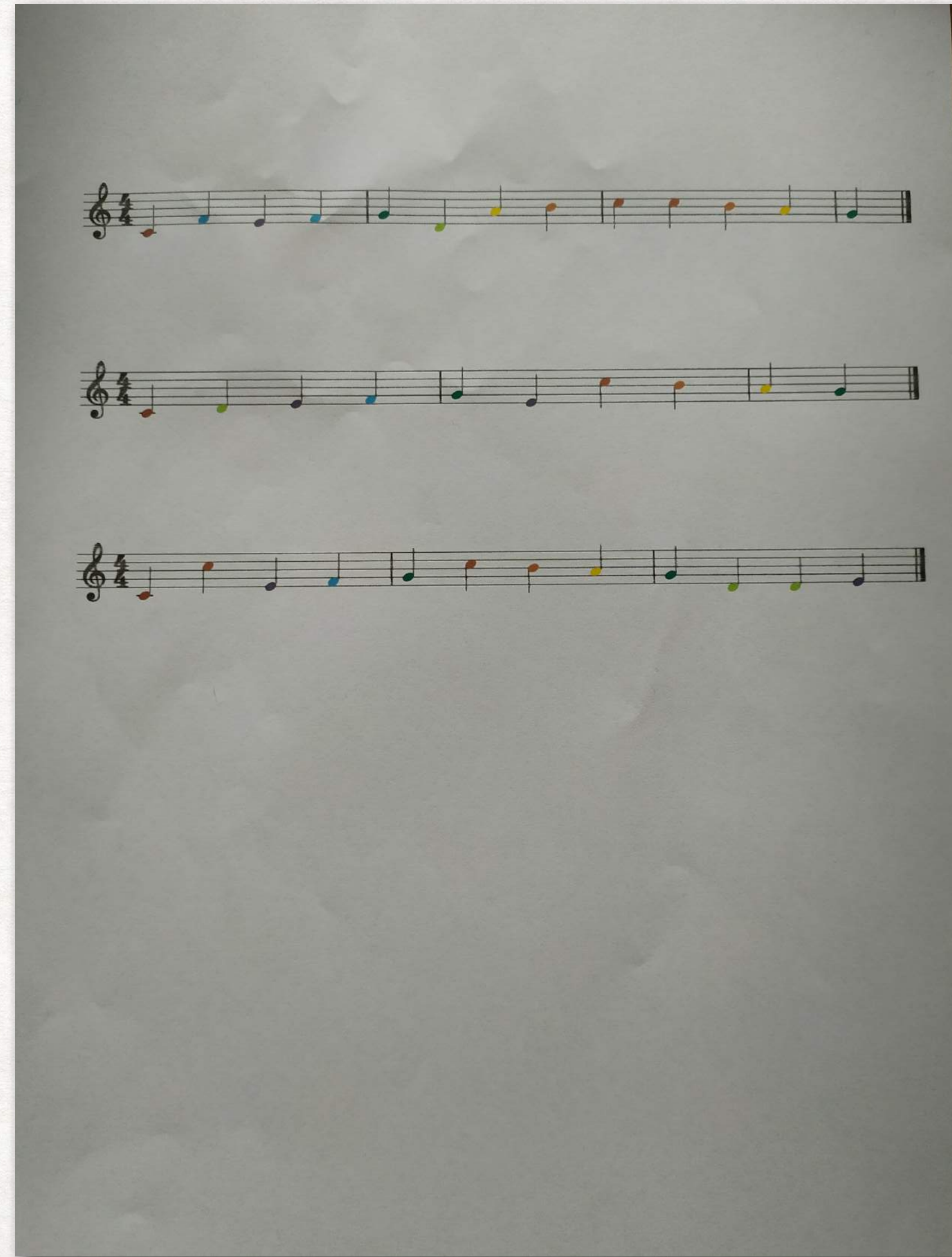


CIEĽ PRÁCE

- Vytvoriť systém na rozpoznanie farebných nôt v notovom zápise a rekonštrukciu notového záznamu s iným farbením

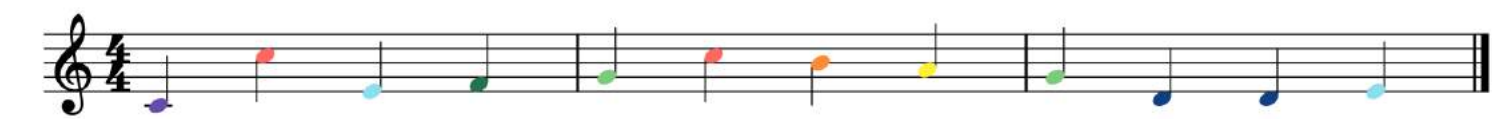
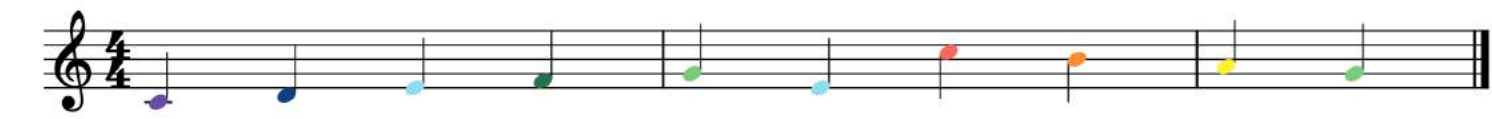
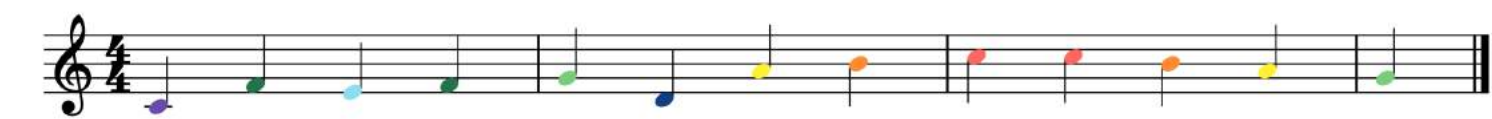
VSTUP

- Fotografia notového záznamu



VÝSTUP

- Nový notový záznam k danému xylofónu

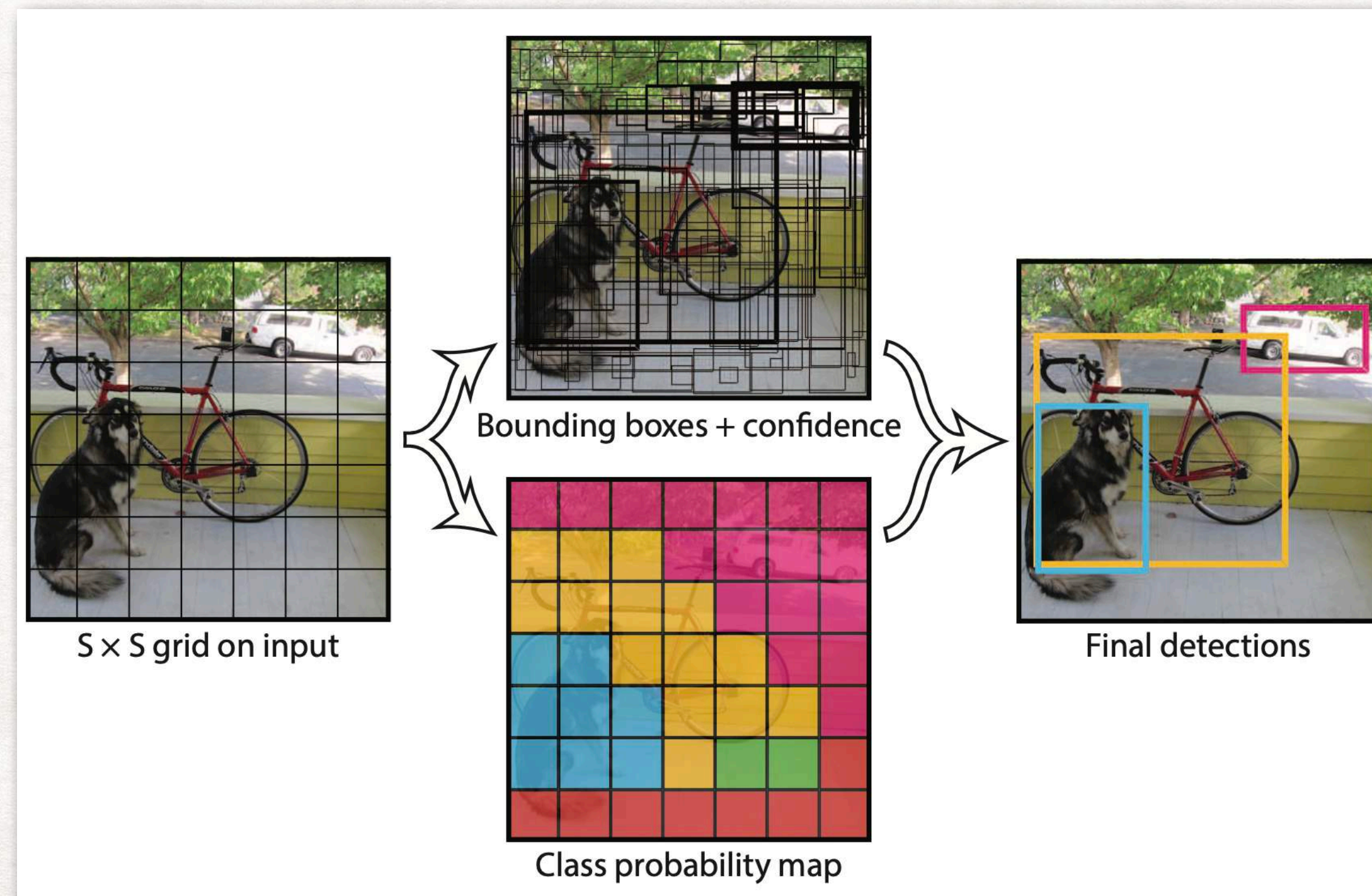


OPTICKÉ ROZPOZNÁVANIE HUDBY (OMR)

- OMR - proces prevodu tlačенých alebo ručne písaných notových záznamov do strojovo čitateľného formátu pomocou spracovania obrazu a techník rozpoznávania vzorov
- Detekcia a rozpoznanie pomocou neurónových sietí (vysoká presnosť, robustnosť na variácie, vieme ich prispôbiť na rôzne problémy)

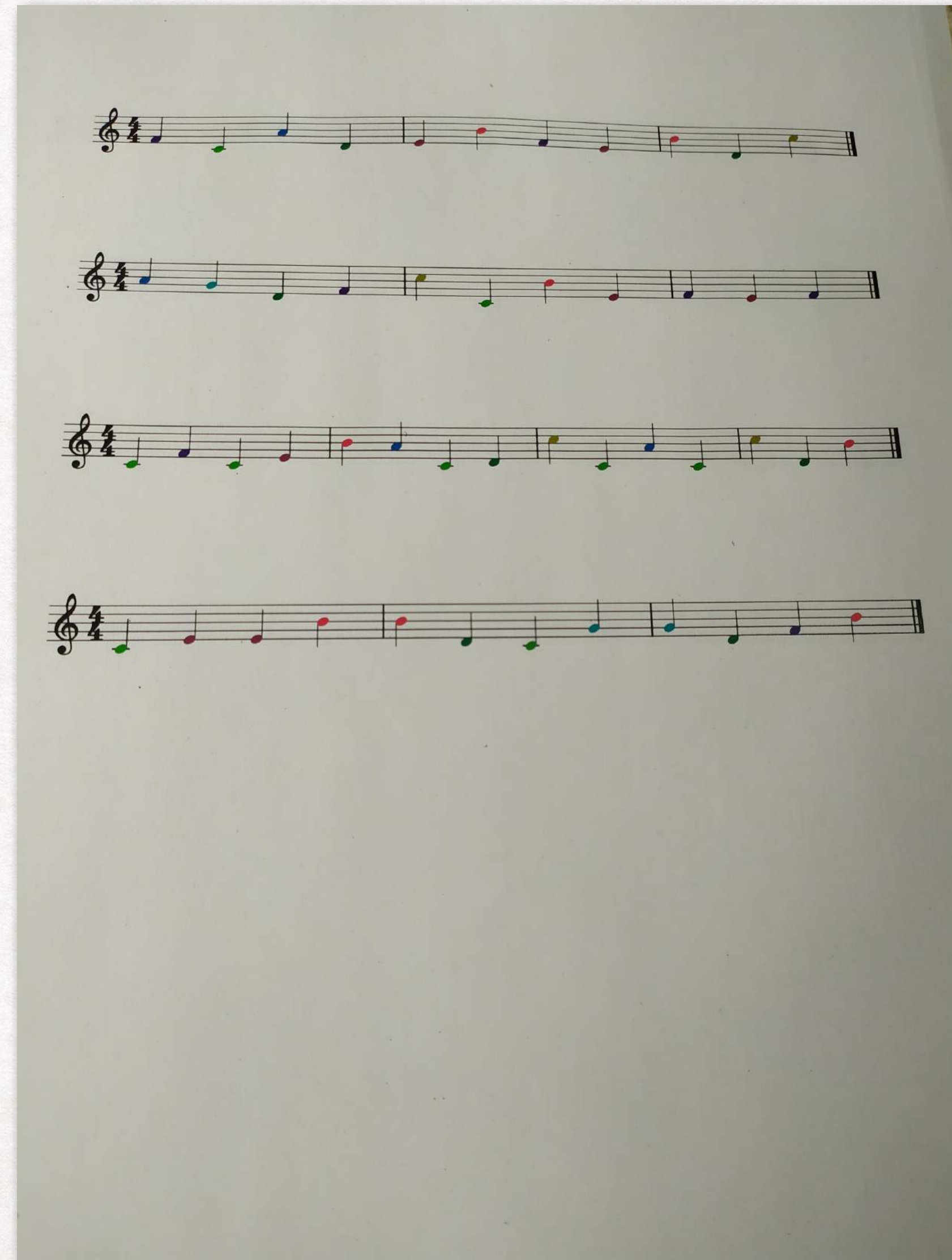
YOLO

- State of art systém detekcie objektov v reálnom čase
- Delí obraz do oblastí a predpovedá ohraničenie a pravdepodobnosť pre každú oblasť



DÁTA

- Ideálny stav: reálne dáta
- Anotovanie reálných fotografií
- **časovo náročné**



DÁTA

- Náhodne generované dáta
- Predpočítané súradnice anotácie

The image shows two musical staves in 4/4 time, each with 16 notes. The notes are represented by colored dots on a five-line staff. Above each staff is a sequence of chord labels in green text. The first staff has labels: G4 D4 C4 A4, E4 G4 D4 B4, E4 A4 C5 C4, C4 C4 A4 F4. The second staff has labels: D4 F4 G4 B4, G4 G4 D4 F4, C4 E4 E4 C5, G4 C5 F4 E4.

G4 D4 C4 A4 E4 G4 D4 B4 E4 A4 C5 C4 C4 C4 A4 F4

D4 F4 G4 B4 G4 G4 D4 F4 C4 E4 E4 C5 G4 C5 F4 E4

TRÉNOVANIE

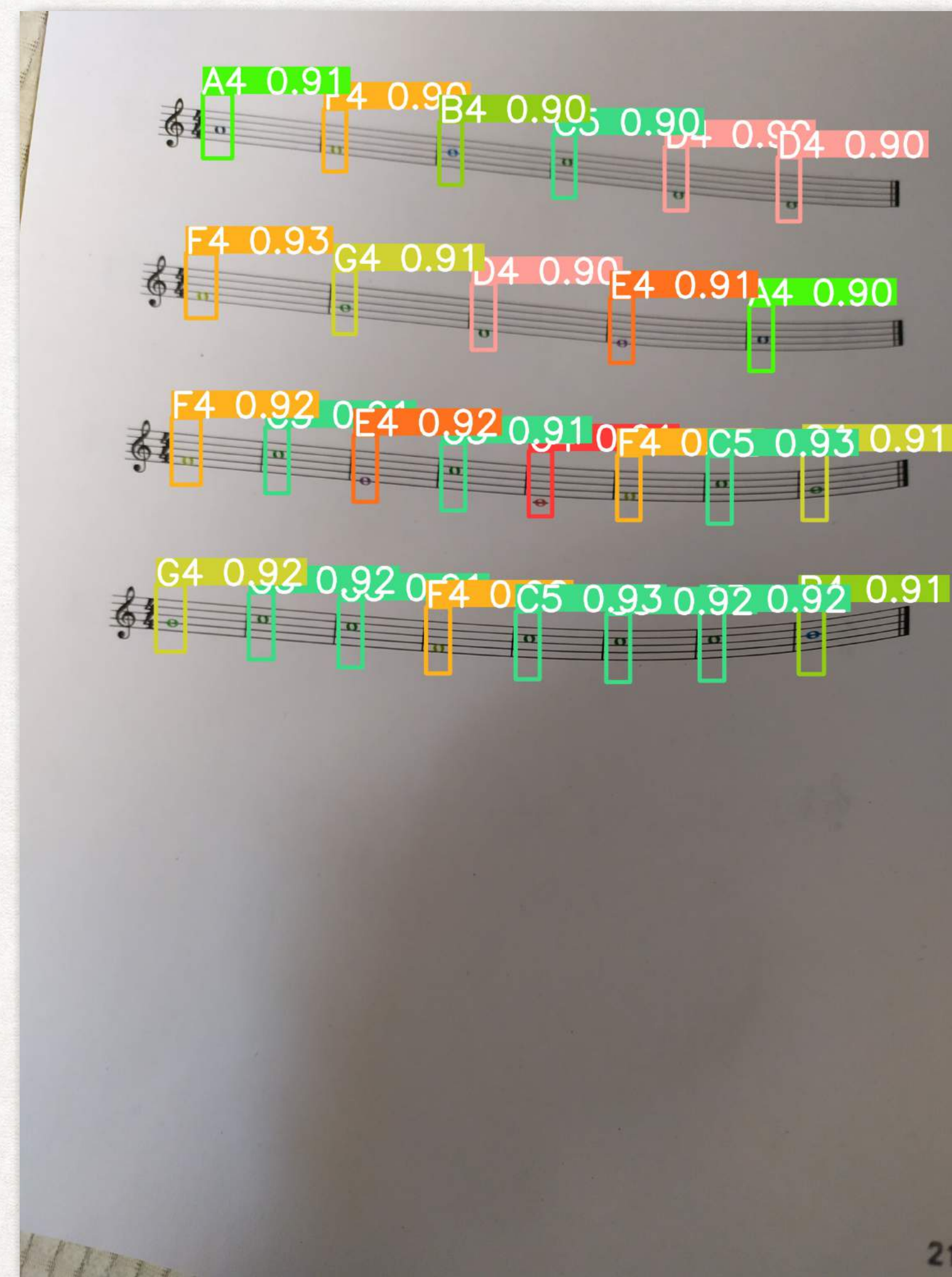
- Modely [yolov8n.pt](#) a [yolov8m.pt](#)
- 2400 trénovacích dát
- 600 validačných dát
- 100 epoch

Model	size (pixels)	mAP ^{val} 50-95	Speed CPU ONNX (ms)	Speed A100 TensorRT (ms)	params (M)	FLOPs (B)
YOLOv8n	640	37.3	80.4	0.99	3.2	8.7
YOLOv8s	640	44.9	128.4	1.20	11.2	28.6
YOLOv8m	640	50.2	234.7	1.83	25.9	78.9
YOLOv8l	640	52.9	375.2	2.39	43.7	165.2
YOLOv8x	640	53.9	479.1	3.53	68.2	257.8

DETEKCIA

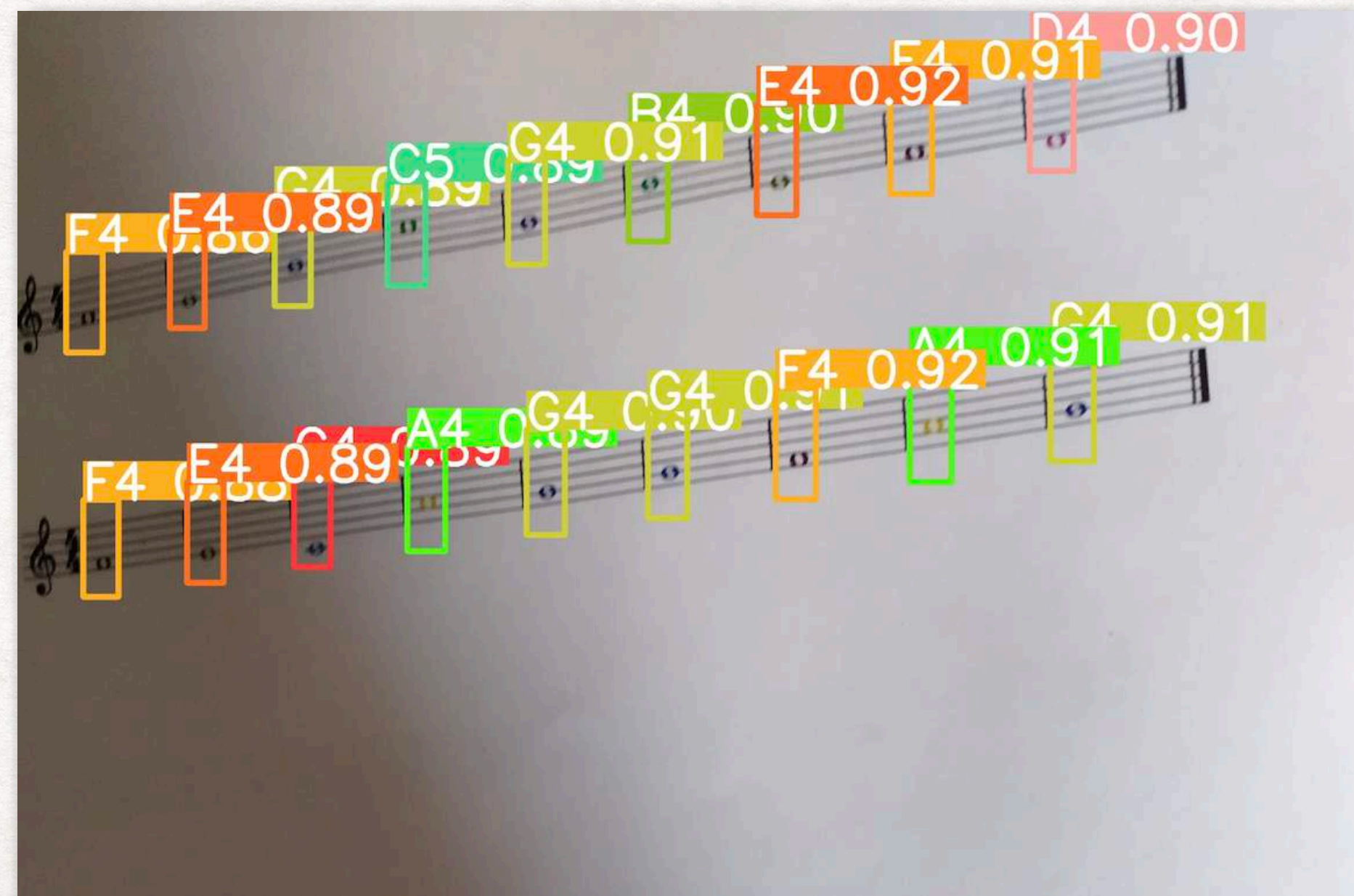
- Trieda, pravdepodobnosť, X, Y, šírka, dĺžka

```
7 0.930466 0.53423077 0.50068986 0.025372233 0.053431276
7 0.9287432 0.7361346 0.37778515 0.025427978 0.053560853
3 0.92518187 0.19051228 0.21675692 0.032318134 0.05107248
3 0.92481333 0.43983617 0.49658814 0.025015015 0.05234146
7 0.92441607 0.6293847 0.5027538 0.025137492 0.05354251
7 0.9243925 0.72802085 0.50135577 0.025602864 0.05451529
2 0.92122656 0.36319548 0.36108845 0.025039673 0.052390777
4 0.92090225 0.15846458 0.47925672 0.030545786 0.05249771
```



REKONŠTRUKCIA

- Poradie nôt
- Usporiadanie z ľava a z hora podľa X,Y môže viesť k zlým výsledkom
- Rozdeliť do skupín podľa riadkov a následne usporiadať

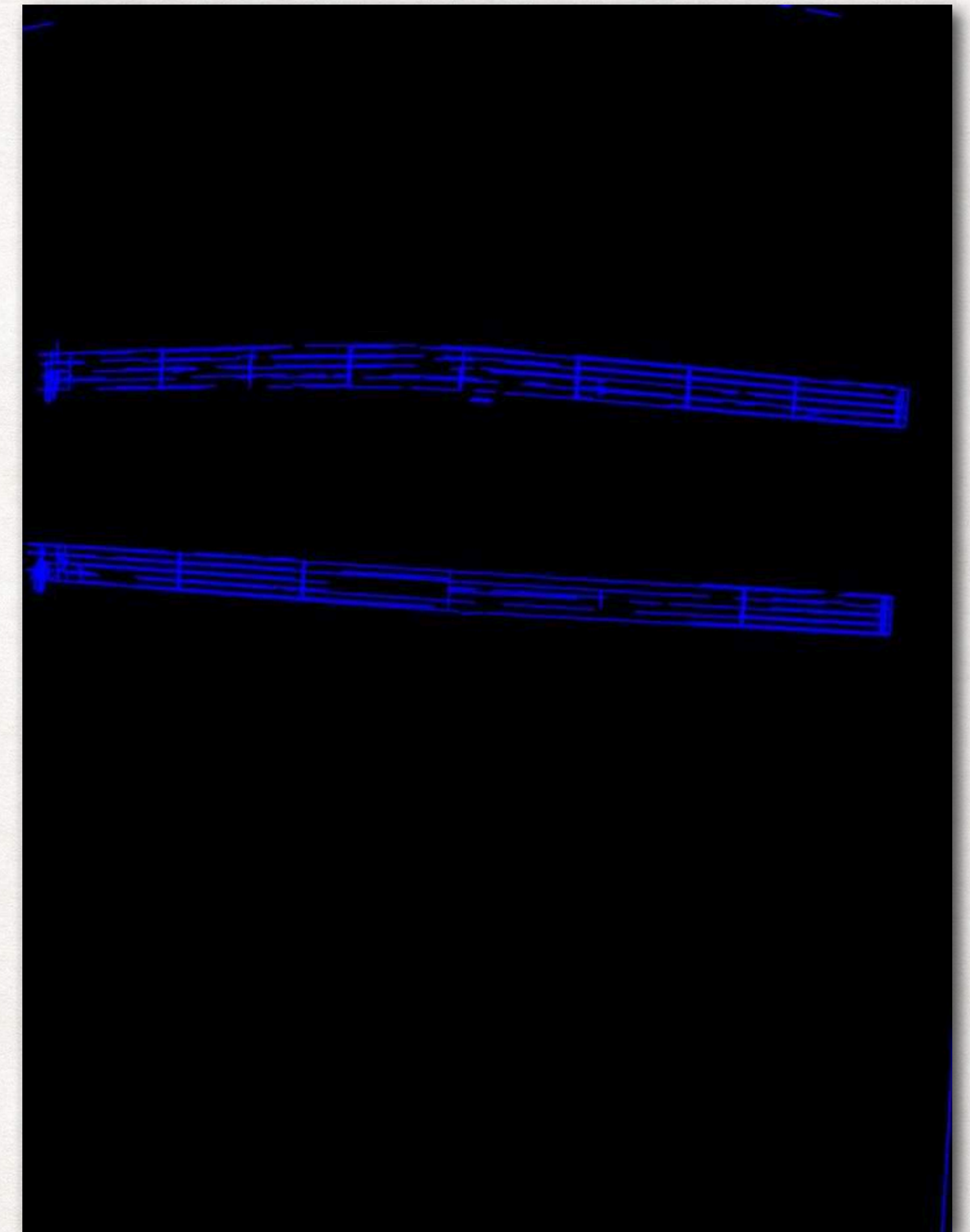
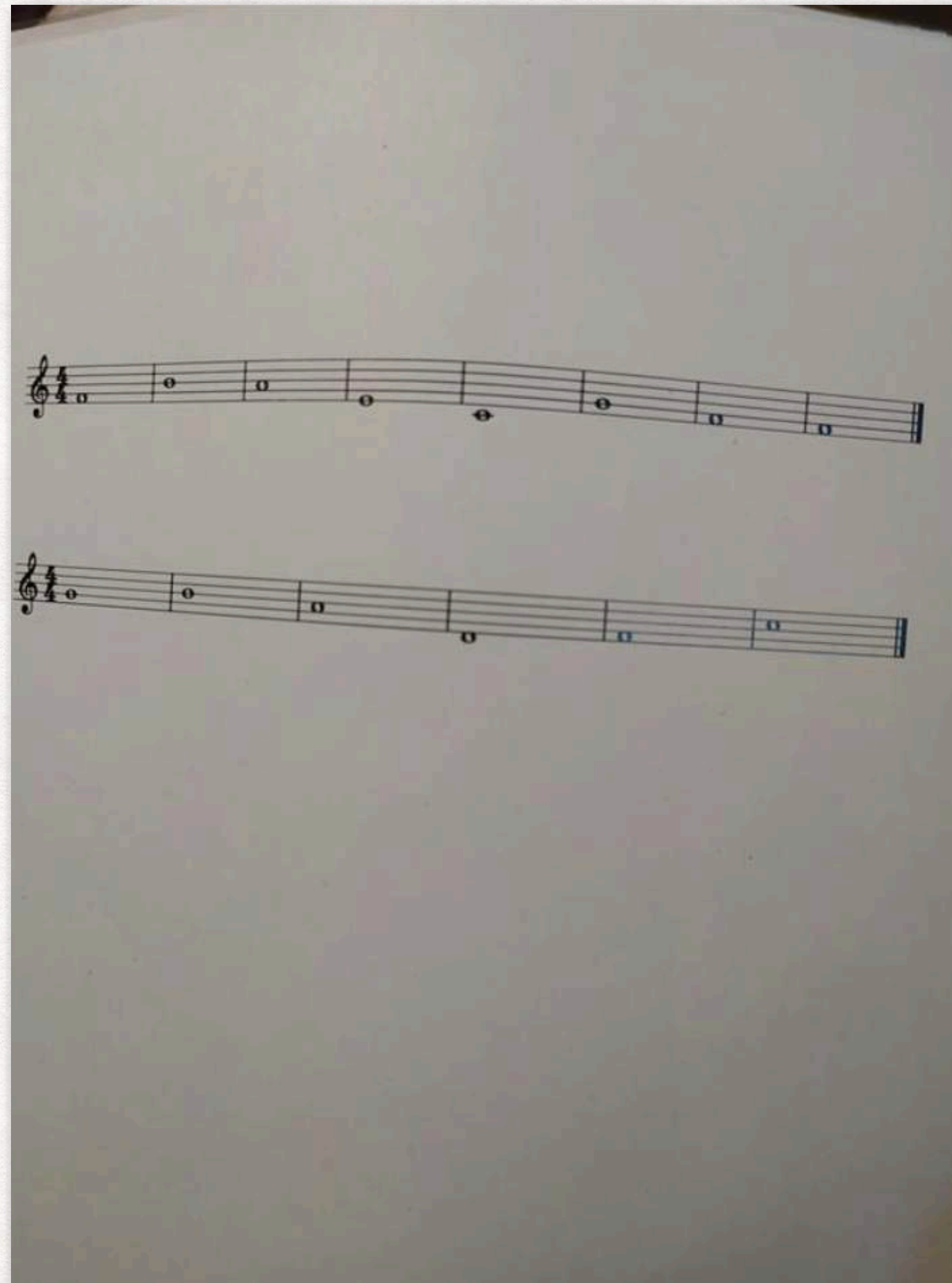


METÓDY RIEŠENIA

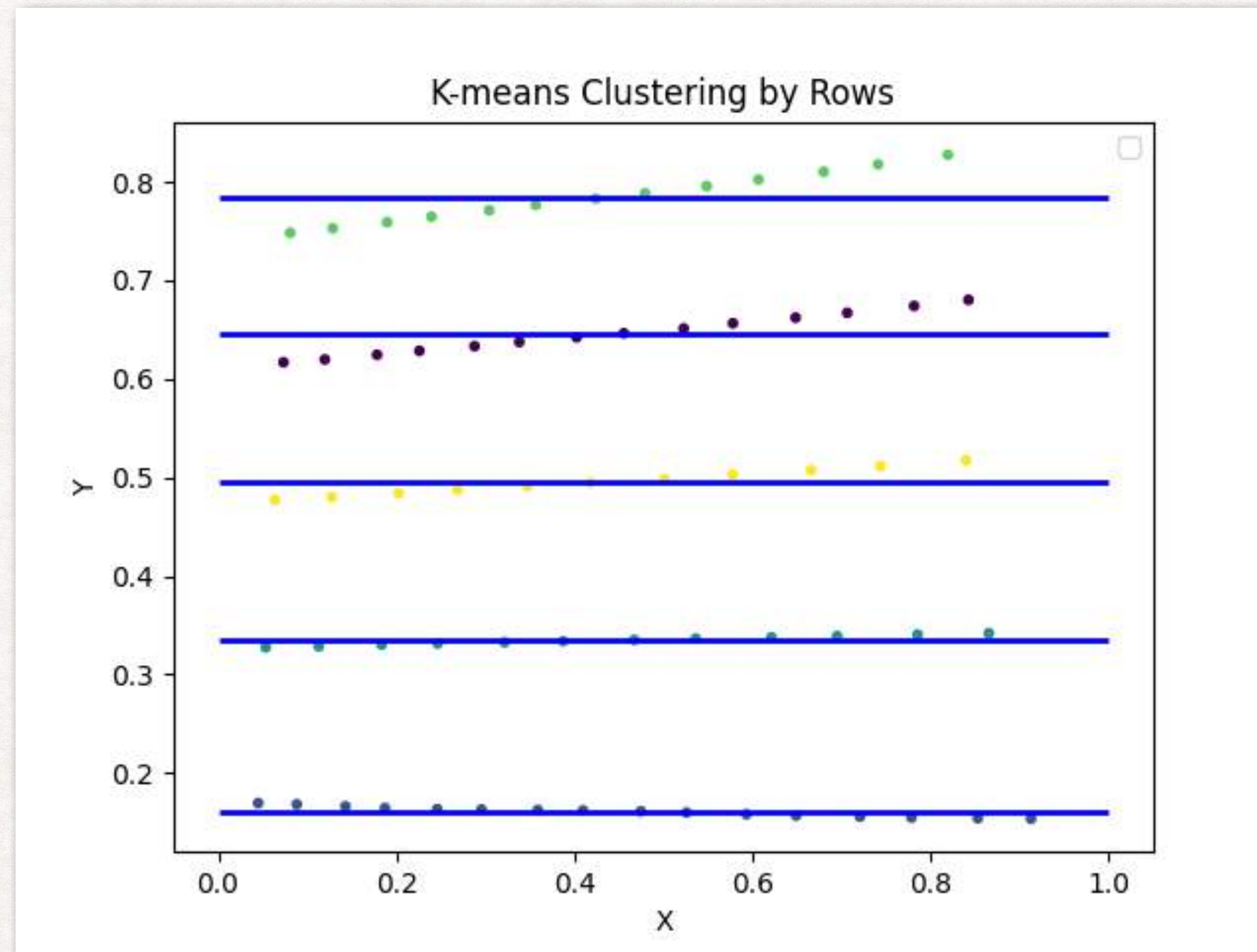
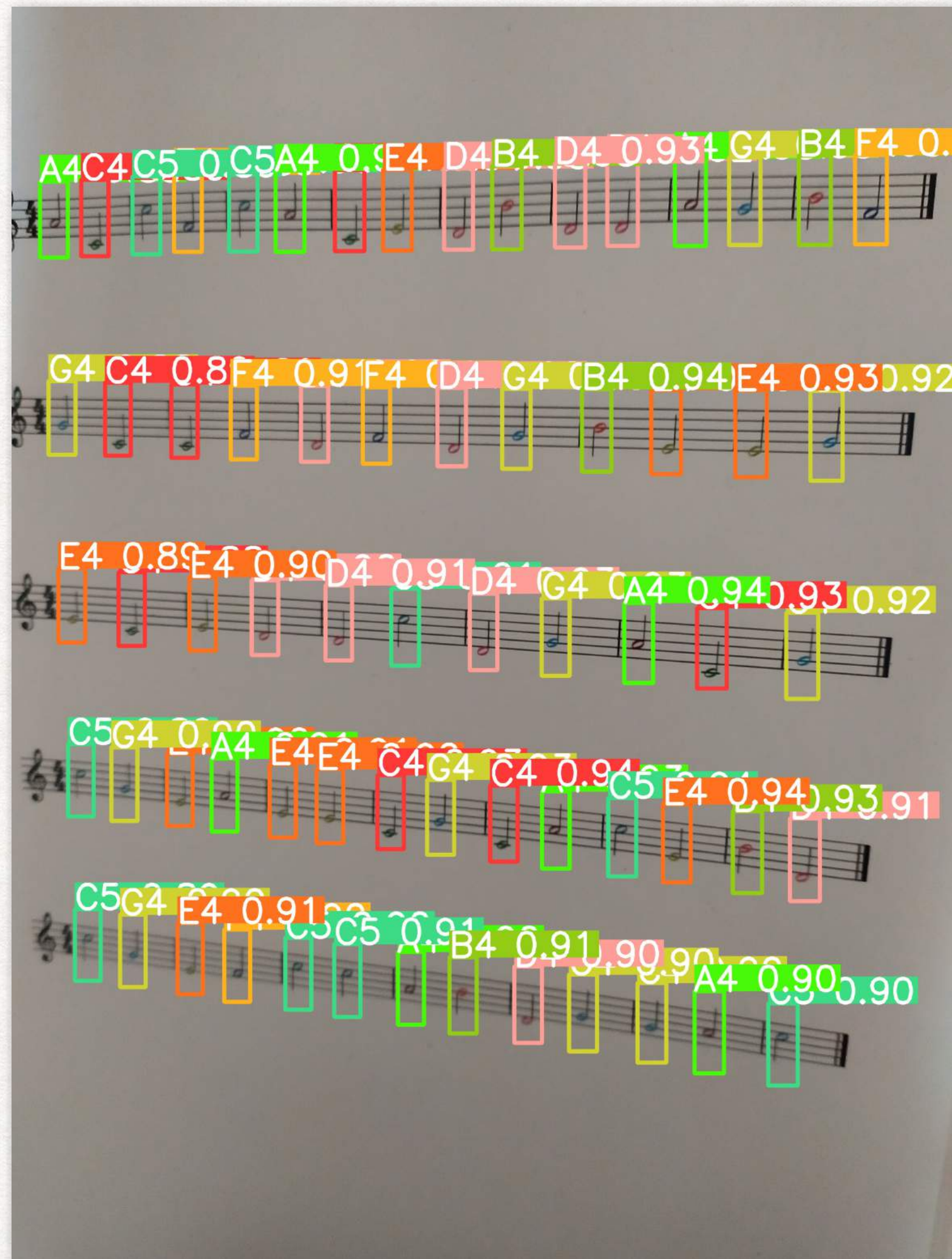
- Houghova transformácia
- K-means Clustering
- Random sample consensus (RANSAC)

HOUGHOVA TRANSFORMÁCIA

- Detekcia čiar
- **Potreba ďalšieho spracovania**

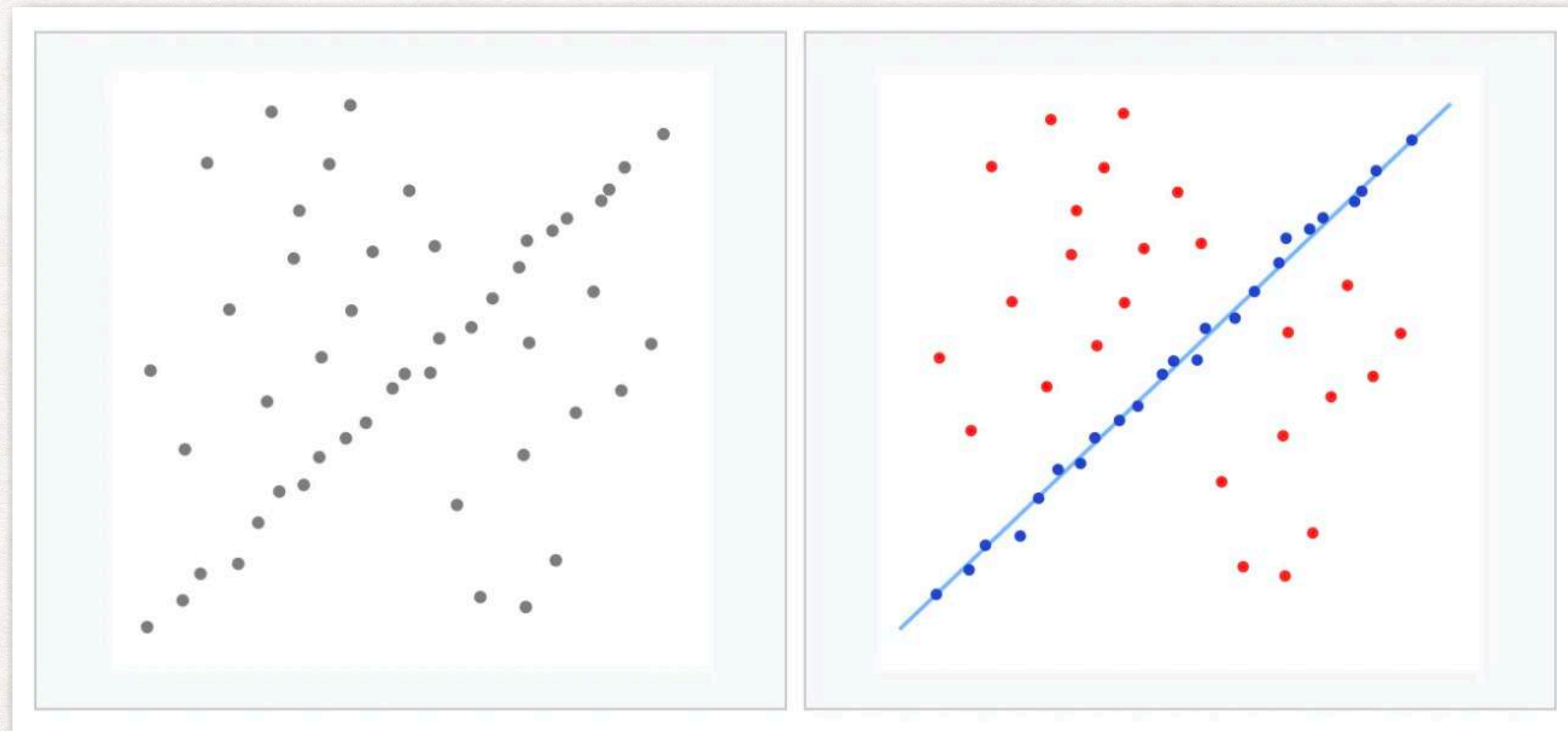


K-MEANS CLUSTERING

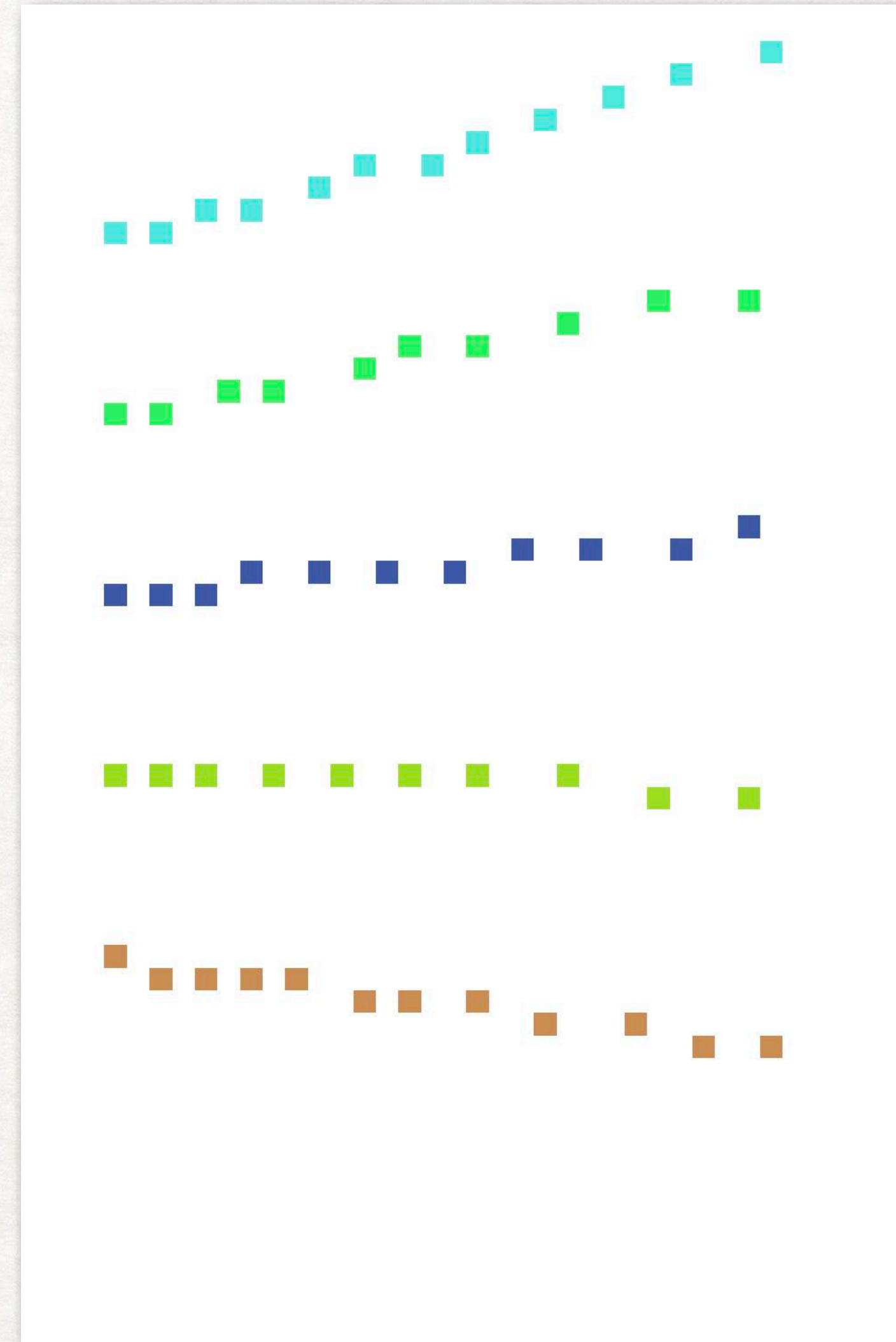
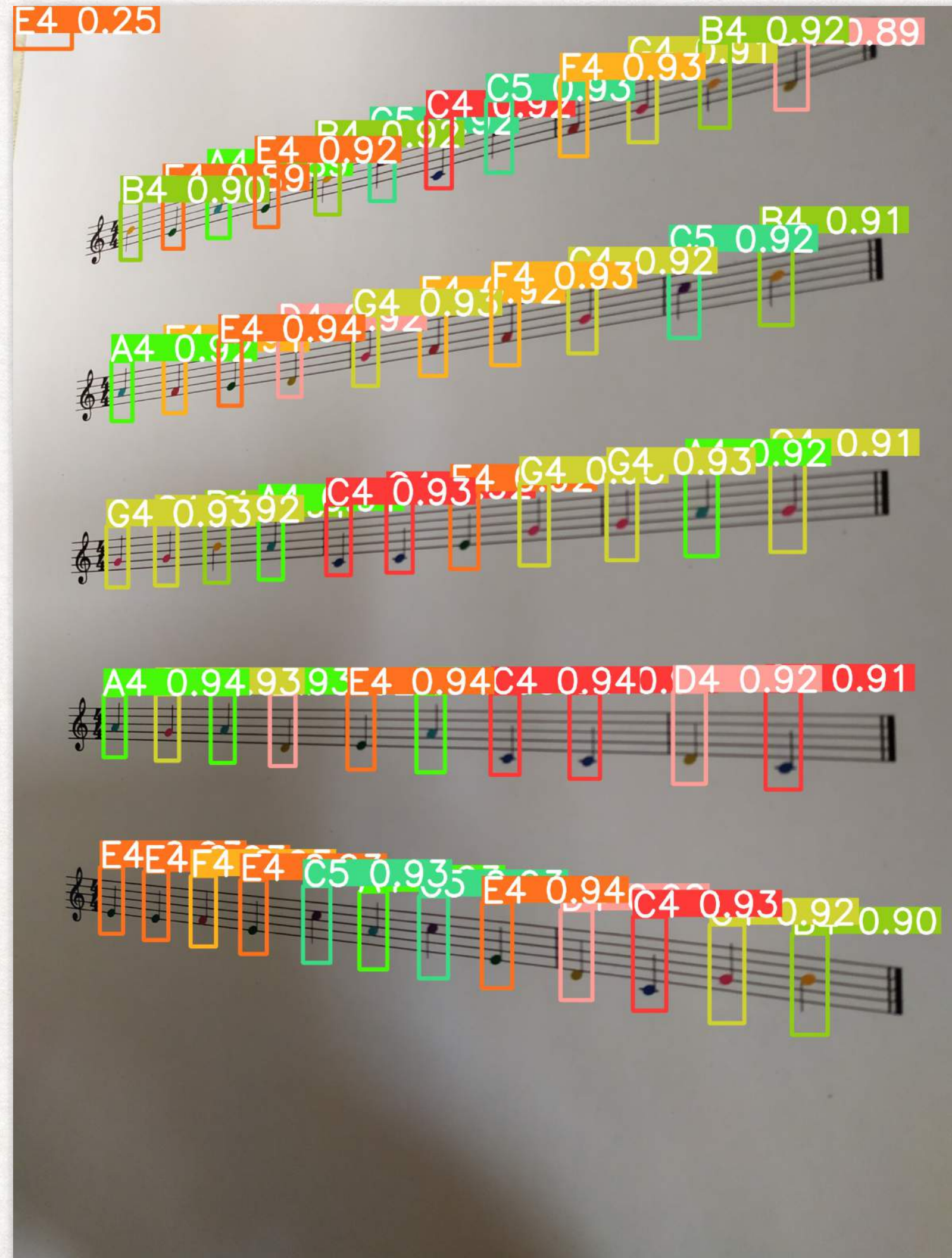


RANSAC

- Iteratívna metóda na odhad parametrov matematického modelu zo súboru pozorovaných údajov, ktorý obsahuje odľahlé hodnoty, pričom odľahlým hodnotám sa nepriznáva žiadny vplyv na hodnoty odhadov.

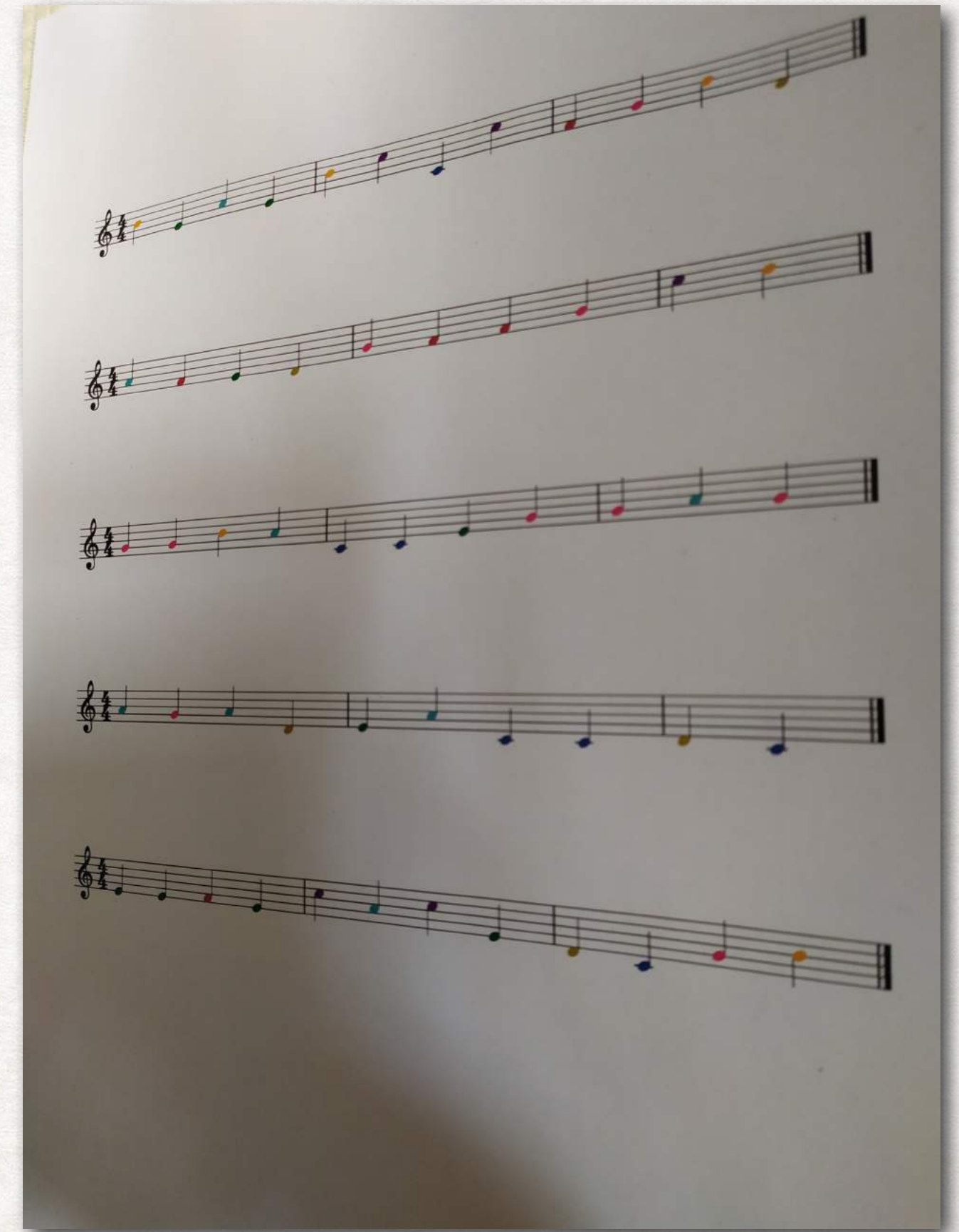
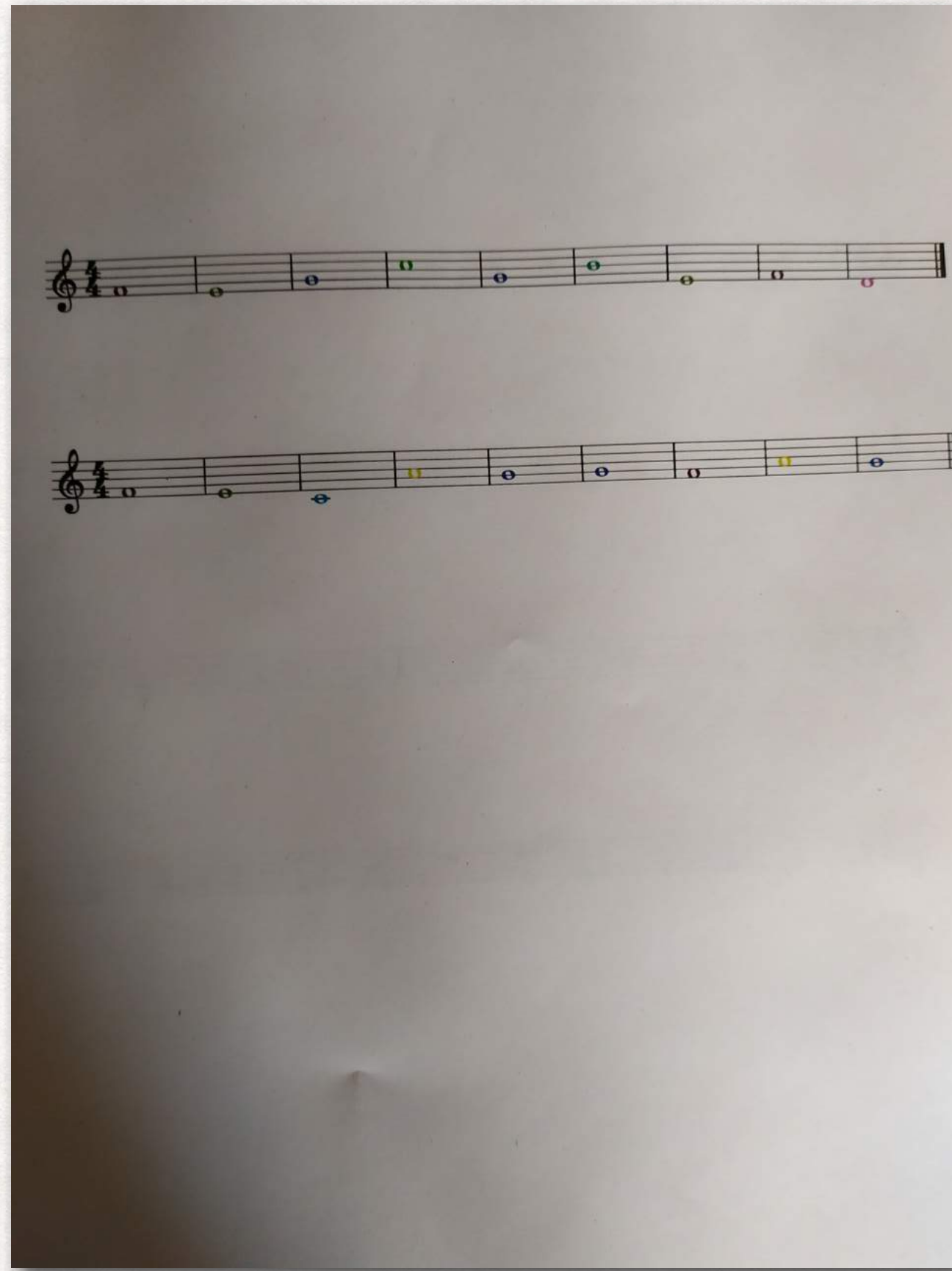
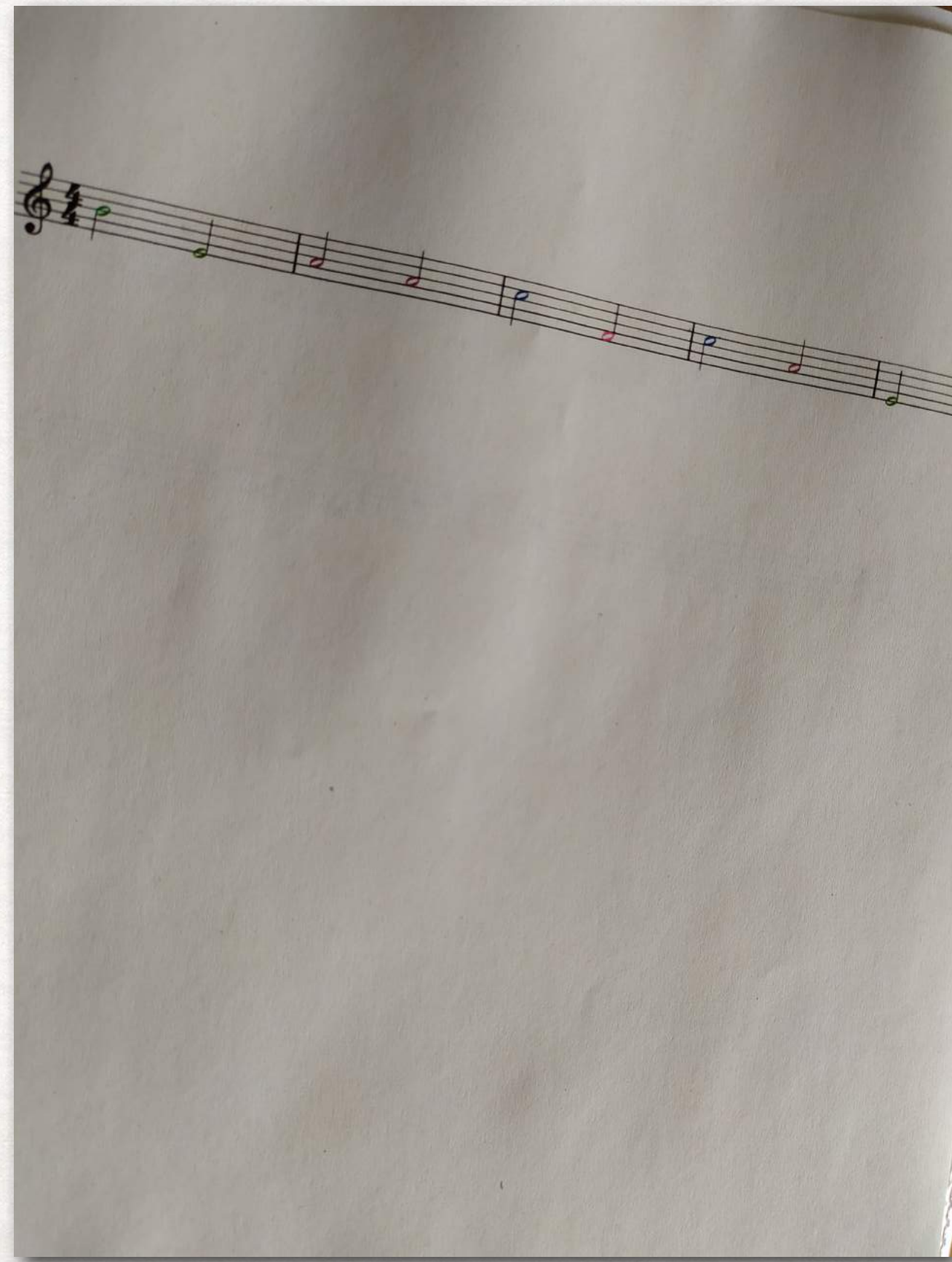


RANSAC



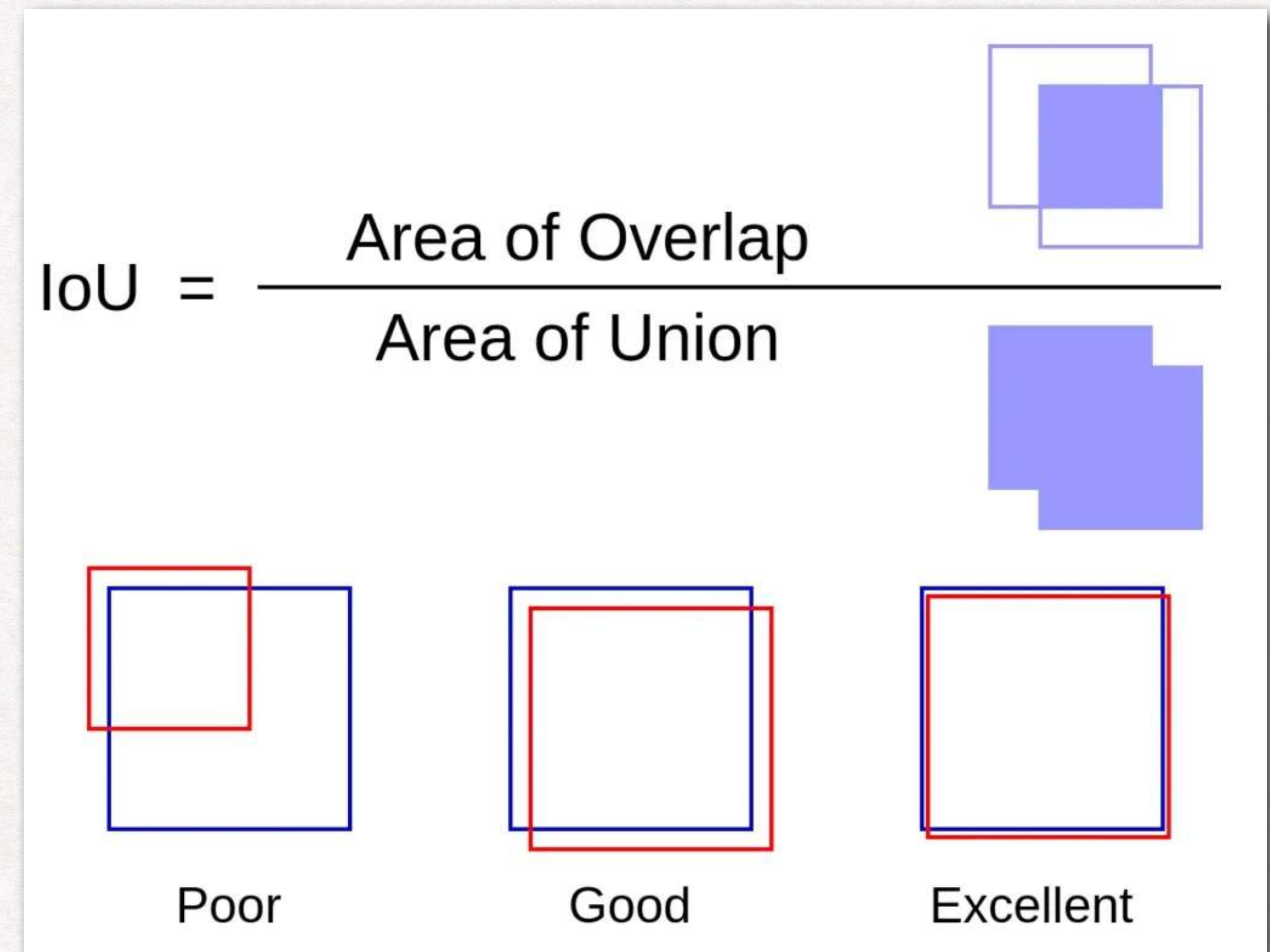
VALIDÁCIA

- 90 fotografií notových záznamov zachytených za rôznych podmienok



VALIDÁCIA

- Intersection over Union (IoU)
- Nemáme anotáciu pre fotografie
- **časovo náročné**



VALIDÁCIA

- Levenshtein algoritmus - meria rozdiel medzi dvoma sekvenciami
- minimálny počet úprav jedného znaku (vloženie, vymazanie alebo nahradenie), ktoré sú potrebné na zmenu jedného slova na druhé.
- Porovnali sme vygenerovanú postupnosť nôh s detekovanou a usporiadanou postupnosťou

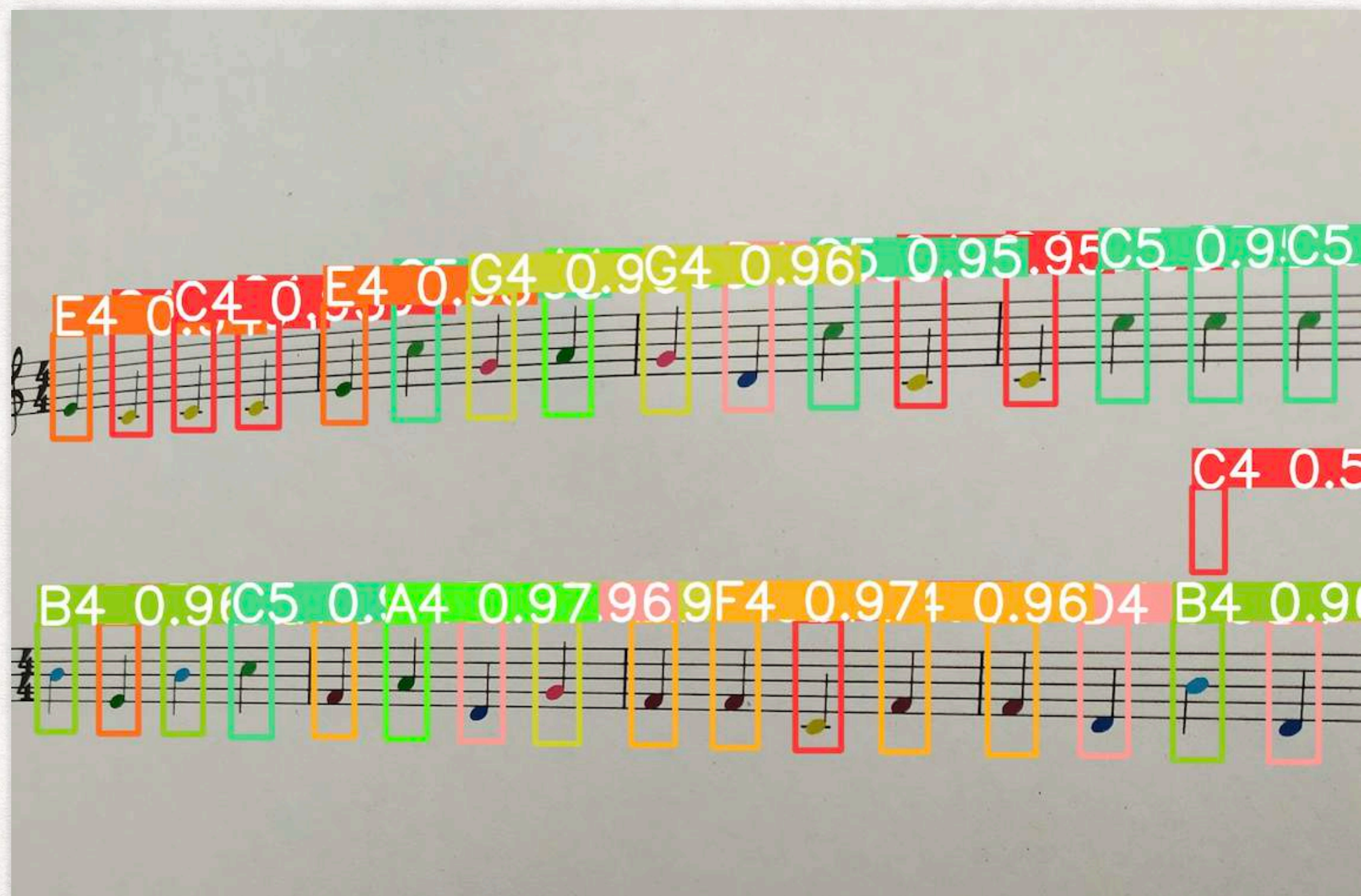
VYHODNOTENIE

- Bez ďalšieho spracovania

Model	K-means	RANSAC
yolov8n.pt	95.41%	95.60%
yolov8m.pt	99.33%	99.66%

VYHODNOTENIE

- RANSAC ignoruje detekcie mimo notovej osnovy



ZLEPŠENIE

- Odstránenie duplicitných detekcií rovnakých objektov s nižšou pravdepodobnosťou

```
6 0.880773 0.4068899 0.43500093 0.031767335 0.06882065
0 0.8729809 0.70904475 0.43139616 0.031286214 0.064492
5 0.8688291 0.42950764 0.18935324 0.026722454 0.055536196
1 0.8664585 0.5525127 0.4329422 0.031199016 0.06679181
5 0.86625355 0.4520293 0.3020805 0.028297607 0.059219148
4 0.86528254 0.13513035 0.30294004 0.029117493 0.060628846
1 0.8633866 0.11416471 0.43857026 0.034480907 0.07132663
5 0.86331797 0.5966866 0.19282287 0.026418254 0.054851778
0 0.86194474 0.26111674 0.4370445 0.032586608 0.0697124
4 0.8602927 0.19532414 0.3024365 0.02831307 0.059648376
2 0.8587961 0.3918324 0.30151927 0.028801676 0.05967987
7 0.85313815 0.8187948 0.3047659 0.028107015 0.05948633
5 0.8498009 0.851877 0.4306311 0.031571615 0.066434115
3 0.84766877 0.3197311 0.30193964 0.028816711 0.059672363
6 0.84535897 0.48326823 0.1917363 0.026222046 0.054988418
6 0.8413088 0.2571501 0.30215785 0.027978394 0.059348054
6 0.84127915 0.6479039 0.19346529 0.026330404 0.056294035
7 0.8400067 0.57352704 0.30188873 0.02708081 0.057106964
0 0.8390752 0.20902486 0.18893017 0.02599764 0.05464862
6 0.834035 0.512878 0.30202124 0.02798755 0.0587041
0 0.83018386 0.76092756 0.30390668 0.027645264 0.058587495
5 0.82699156 0.69836557 0.1947792 0.026951741 0.05568666
0 0.81875837 0.64397174 0.30202743 0.02714205 0.056675598
2 0.81055635 0.80819744 0.20088671 0.026421793 0.055889342
1 0.7971984 0.85744697 0.20311283 0.025624268 0.05368599
4 0.78540576 0.26085678 0.18789558 0.025635254 0.05412982
6 0.7844393 0.8882413 0.30558807 0.028998697 0.060452666
0 0.77001894 0.70295507 0.3025678 0.02700244 0.05751068
1 0.7679472 0.37748805 0.18732496 0.027265096 0.055490494
2 0.74789196 0.7492729 0.19756629 0.026307454 0.05565759
2 0.74583435 0.3131148 0.18692794 0.026252503 0.054979905
2 0.7383333 0.53439337 0.1917457 0.026834553 0.056173585
4 0.6303154 0.15456524 0.19106124 0.025227234 0.053412
3 0.6028812 0.3129932 0.18761928 0.026338927 0.05534796
3 0.5053517 0.74902767 0.19804987 0.026803955 0.056479752
1 0.45483384 0.8078605 0.2004712 0.025727782 0.05521077
3 0.41924316 0.1546069 0.19061708 0.025572479 0.053912964
2 0.38757303 0.85758543 0.2033142 0.02649764 0.0547478
```


VYHODNOTENIE PO ZLEPŠENÍ

Model	K-means	RANSAC
yolov8n.pt	98.43%	98.66%
yolov8m.pt	99.67%	100%

VYSOKÁ ÚSPEŠNOST

- Jednoduchý podproblém
- Neprekrývajú sa hudobné symboly
- YOLO dosahuje vysokú úspešnosť aj pri zložitých detekciach
- Trénovacie a testovacie dáta majú rovnakú štruktúru

ZHRNUTIE

- Navrhli sme a otestovali systém na detekciu a rekonštrukciu farebného notového záznamu
- Tvorba dát
- Trénovanie modelu
- Usporiadanie detekovaných objektov
- Vyhodnotenie modelu

REKONŠTRUKCIA



+



A musical score with three staves. The notes are colored and have MIDI pitch data above them. The data includes notes like C4, F4, E4, B4, A4, D4 and their corresponding frequencies and intervals.

Three horizontal rows of colored dots: yellow, blue, and green.



+

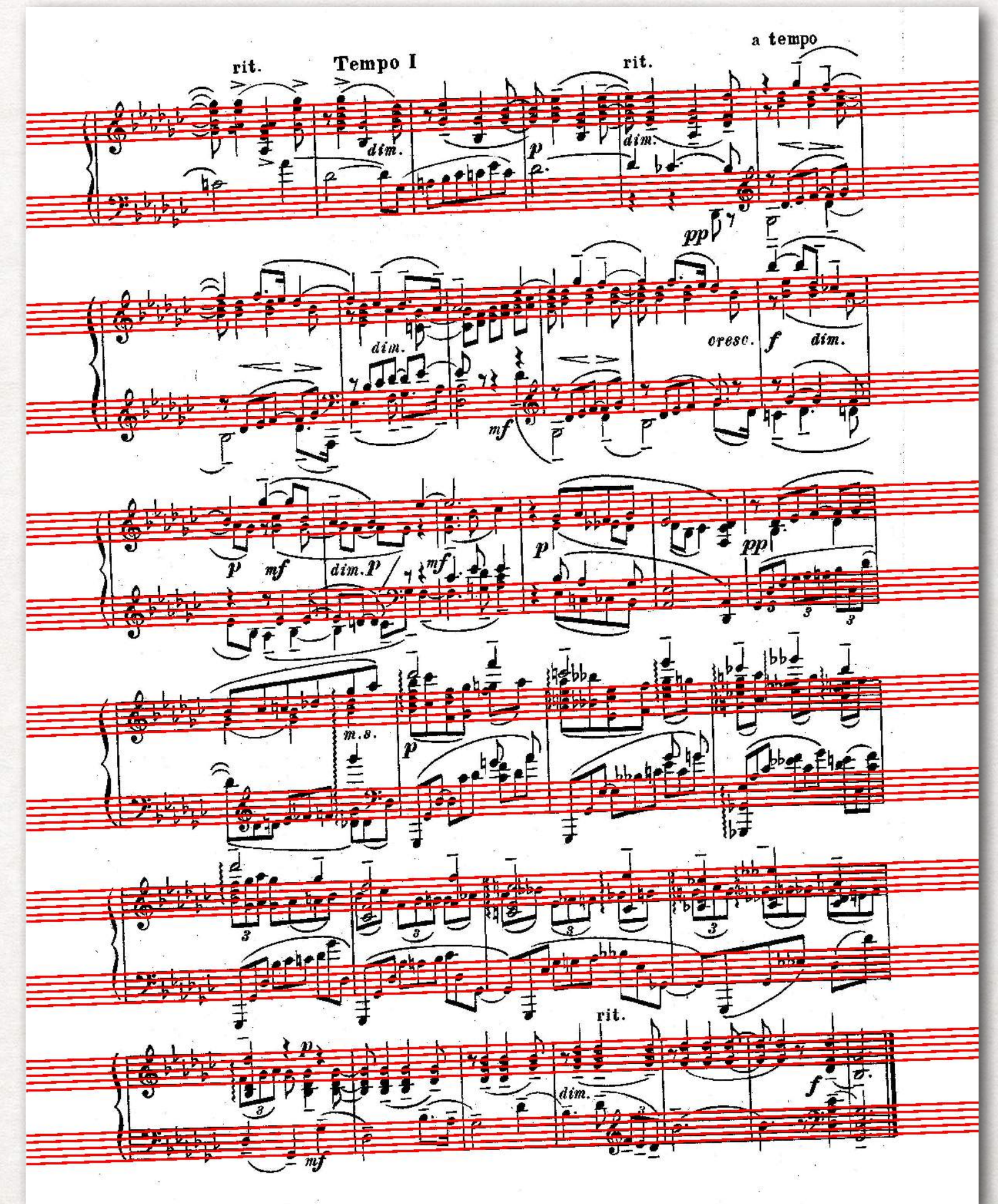
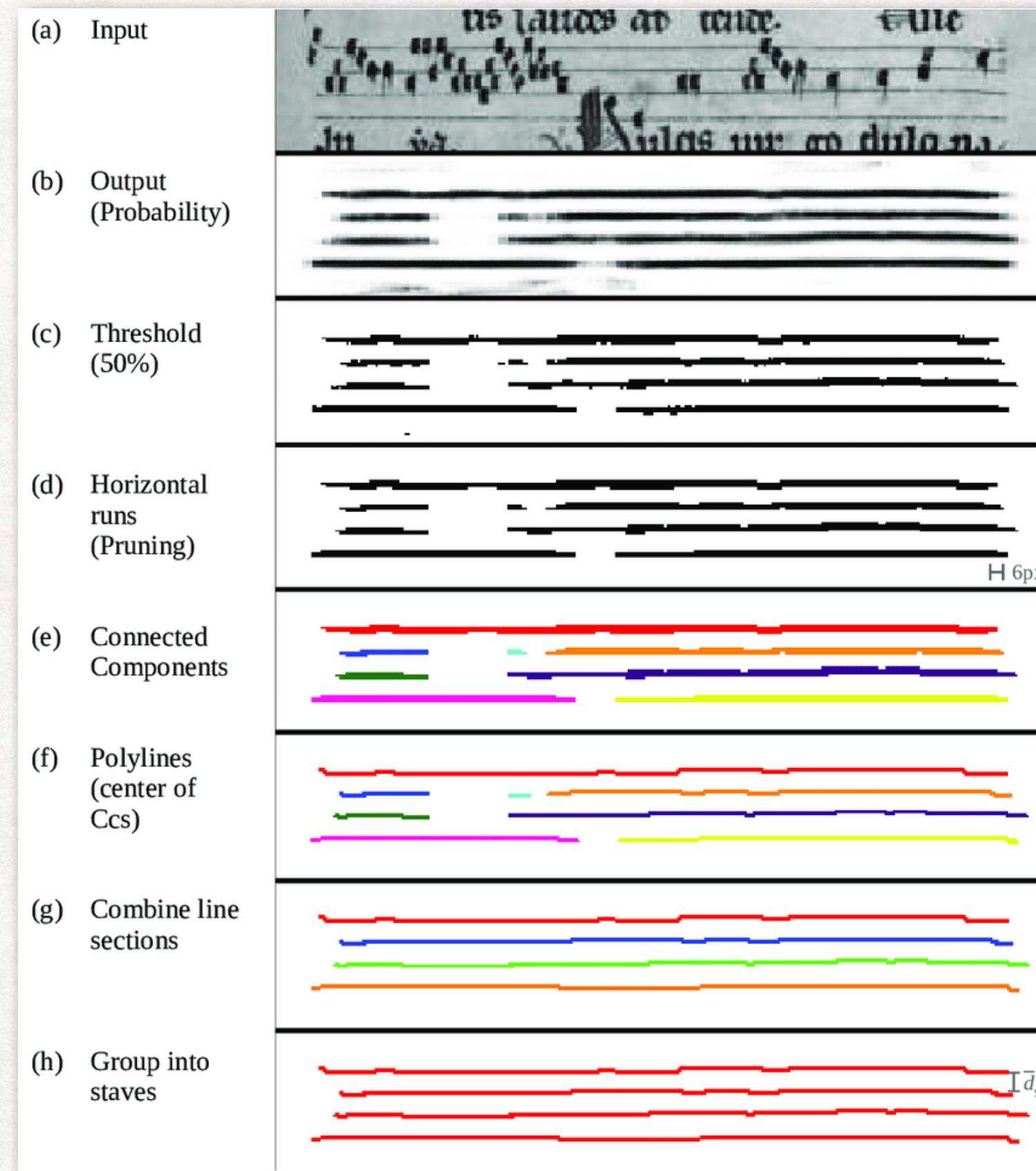
A musical score with three staves. The notes are colored to match the xylophone bars.

A musical score with three staves. The notes are colored to match the xylophone bars.

ĎAKUJEM ZA POZORNOST

POČET SKUPÍN

- Notová osnova sa skladá z 5 čiar
- Detekcia čiar
- Houghova transformácia
- Skew projection

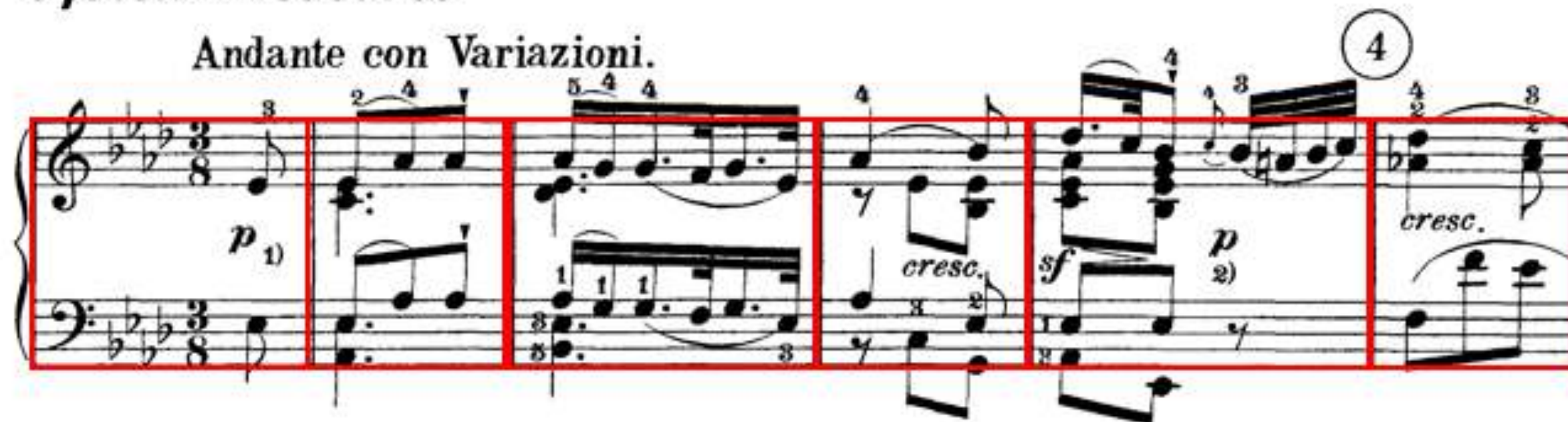


POČET SKUPÍN

- Detekčný problém
- Dotrénovať model

System Measures

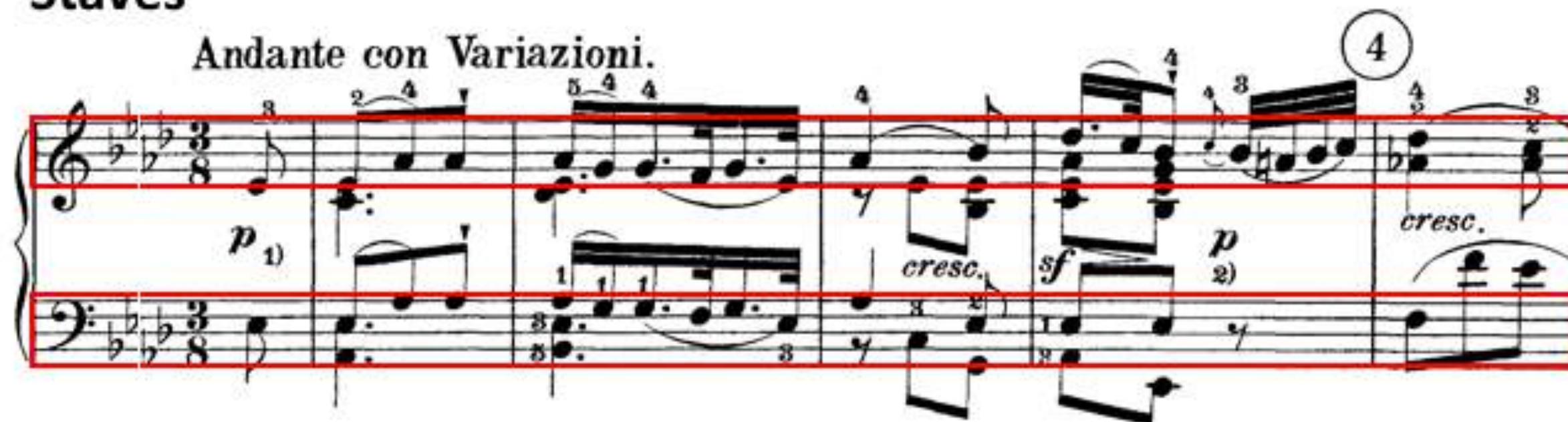
Andante con Variazioni.



A musical score for piano in 3/8 time, titled "Andante con Variazioni." The score consists of two staves. The first staff is marked with a circled '4' above the final measure. The second staff is marked with a circled '4' above the final measure. The score is divided into six measures by vertical red lines. The first measure is marked with a circled '8' above the staff and a circled '4' below the staff. The second measure is marked with a circled '2' above the staff and a circled '4' below the staff. The third measure is marked with a circled '5' above the staff and a circled '4' below the staff. The fourth measure is marked with a circled '4' above the staff and a circled '4' below the staff. The fifth measure is marked with a circled '4' above the staff and a circled '4' below the staff. The sixth measure is marked with a circled '4' above the staff and a circled '4' below the staff. The score includes dynamic markings: *p*₁ in the first measure, *cresc.* in the fourth measure, *sf* in the fifth measure, *p*₂ in the sixth measure, and *cresc.* in the seventh measure. The score also includes various musical notations such as slurs, ties, and fingerings.

Staves

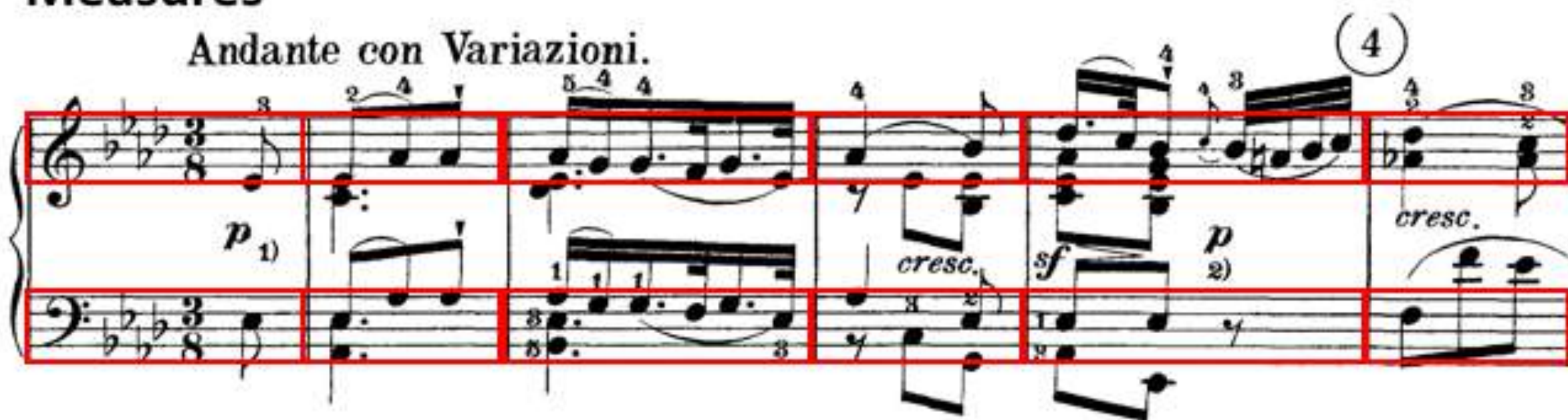
Andante con Variazioni.



A musical score for piano in 3/8 time, titled "Andante con Variazioni." The score consists of two staves. The first staff is marked with a circled '4' above the final measure. The second staff is marked with a circled '4' above the final measure. The score is divided into six measures by vertical red lines. The first measure is marked with a circled '8' above the staff and a circled '4' below the staff. The second measure is marked with a circled '2' above the staff and a circled '4' below the staff. The third measure is marked with a circled '5' above the staff and a circled '4' below the staff. The fourth measure is marked with a circled '4' above the staff and a circled '4' below the staff. The fifth measure is marked with a circled '4' above the staff and a circled '4' below the staff. The sixth measure is marked with a circled '4' above the staff and a circled '4' below the staff. The score includes dynamic markings: *p*₁ in the first measure, *cresc.* in the fourth measure, *sf* in the fifth measure, *p*₂ in the sixth measure, and *cresc.* in the seventh measure. The score also includes various musical notations such as slurs, ties, and fingerings.

Measures

Andante con Variazioni.



A musical score for piano in 3/8 time, titled "Andante con Variazioni." The score consists of two staves. The first staff is marked with a circled '4' above the final measure. The second staff is marked with a circled '4' above the final measure. The score is divided into six measures by vertical red lines. The first measure is marked with a circled '8' above the staff and a circled '4' below the staff. The second measure is marked with a circled '2' above the staff and a circled '4' below the staff. The third measure is marked with a circled '5' above the staff and a circled '4' below the staff. The fourth measure is marked with a circled '4' above the staff and a circled '4' below the staff. The fifth measure is marked with a circled '4' above the staff and a circled '4' below the staff. The sixth measure is marked with a circled '4' above the staff and a circled '4' below the staff. The score includes dynamic markings: *p*₁ in the first measure, *cresc.* in the fourth measure, *sf* in the fifth measure, *p*₂ in the sixth measure, and *cresc.* in the seventh measure. The score also includes various musical notations such as slurs, ties, and fingerings.

MUSICXML.PNG FORMAT

- Kapitola 4.1.3
- Podporované formáty na vykreslenie notového záznamu (musicxml, text, lilypond, lily.png, lily.pdf alebo musicxml.png)
- Vstupný parameter pre funkciu **show(format)**

RÔZNE FONTY

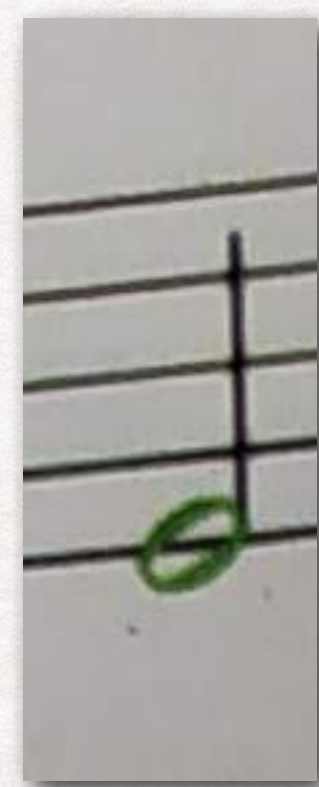
- Netestovali sme
- Stačí rozšíriť tréningovú množinu



NOTY RÔZNEJ DĹŽKY

- Rozšíriť množinu tried
- Detekovať objekty s výškou a v rámci objektu určiť vlastnosť

-  - celá
-  - pólóvá
-  - štvrtinová
-  - osminová
-  - šestnástinová



Annotations for the notes in the staff:

- Notes 1-4: B4 0.95, E4 0.95, B4 0.96, C5 0.96
- Notes 5-8: F4 0.96, G4 0.96, A4 0.97, B4 0.96
- Notes 9-12: C4 0.95, C4 0.96, E4 0.96, G4 0.96
- Notes 13-16: A4 0.96, A4 0.96, E4 0.96, C4 0.95, D4 0.96, D4 0.96

Additional annotations for notes 17-20 (from a second staff):

- Notes 17-20: E4 0.96, E4 0.96, F4 0.96, A4 0.96, C4 0.96, C4 0.96, A4 0.96, E4 0.96