



EÖTVÖS LORÁND TUDOMÁNYEGYETEM

INFORMATIKAI KAR

PROGRAMOZÁSELMÉLET ÉS SZOFTVERTECHNOLÓGIAI
TANSZÉK

Automatikus zenei hangszerfelismerés többszólamú zenében mély neuronhálók segítségével

Témavezető:

Gombos Gergő

Adjunktus, PhD

Szerző:

Hamrák János

programtervező informatikus MSc

Budapest, 2020

Az eredeti szakdolgozati / diplomamunka témabjelentő helye.

Tartalomjegyzék

1. Bevezetés	3
1.1. Motiváció	3
1.2. A dolgozat felépítése	3
1.3. Kapcsolódó munkák	4
2. Elméleti háttér	5
2.1. Zene és reprezentációi	5
2.1.1. Zene fogalma, tulajdonságai	5
2.1.2. Audio reprezentációk	5
2.1.3. Musical Information Retrieval (MIR)	5
2.2. Machine learning, Deep Learning	5
2.2.1. Machine Learning	5
2.2.2. Deep Learning	6
3. Módszertan	7
3.1. Előfeldolgozás	7
3.2. Architektúra	7
3.3. Megvalósítás	7
3.3.1. Könyvtárak	7
4. Adathalmaz	8
4.1. Kiválasztási szempontok	8
4.2. OpenMIC	8
4.3. MedleyDB	8
4.4. Egyebek	8
5. Kísérletek, eredmények	9

5.1. Mérőszámok	9
5.2. Eredmények	10
5.2.1. próba1	11
5.2.2. próba2	11
5.2.3. próba3	11
5.2.4. próba4	11
5.3. Összehasonlítás	11
5.3.1. saját próbálkozásaim	11
5.3.2. más munkákkal (hagyományos ML)	11
6. Összegzés, kitekintés	12
Irodalomjegyzék	13

1. fejezet

Bevezetés

1.1. Motiváció

Az elmúlt években egyre felkapottabb kutatási területté váltak a mély neuron-hálós rendszerek. Ezen belül a zenei információk kinyerésének kutatása is jelentős teret hódított magának.

1.2. A dolgozat felépítése

Dolgozatomban tehát az eddigi kapcsolódó kutatásokat, illetve saját munkám eredményét dolgozom fel. A következő alfejezetben felsorolom az általam relevánsnak tartott, a State-of-the-Art-hoz vezető kutatásokat.

A második fejezetben betekintést adok a téma elméleti hátterébe. Először kifejtem a zenével kapcsolatos főbb fogalmakat, bemutatom fontosabb tulajdonságait, reprezentációit. Kitérek a kapcsolódó kutatási terület bemutatására is. Ezután bevezetem a gépi tanulás és a mély tanulás fogalmát.

A harmadik fejezetben a módszertanról ejtek szót. Itt kifejtésre kerülnek az adat előfeldolgozási módszerei, az általam bemutott mély tanulási architektúrák, illetve ezek megvalósításai. A negyedik fejezet az adathalmazokról fog szólni. Itt először felsorolom az adathalmazok kiválasztásának szempontjait, majd minden felhasznált adathalmaznak leírom a főbb jellemzőit.

Az ötödik fejezetben részletezem az általam végzett kísérleteket és ezek eredményeit. Eután ezeket összevetem egymással, illetve a releváns State-of-the-Art kutatá-

sokkal. A hatodik fejezetben összegzem a leírtakat és kutatásom továbbgondolására felvázolok néhány ötletet.

1.3. Kapcsolódó munkák

Cite author Li, Qian és Wang and cite [2].

2. fejezet

Elméleti háttér

felsorolni hogy miket kell leírni nagyjából

2.1. Zene és reprezentációi

audio, reprezentációk, "deep learning in mir" cikk és egyéb forrásokból

2.1.1. Zene fogalma, tulajdonságai

bevezető a zene hétköznapi fogalmáról

2.1.2. Audio reprezentációk

a hangokról kicsit mélyebben, reprezentációk, stb.

2.1.3. Musical Information Retrieval (MIR)

IR, MIR, time varying, static problémák, etc.

2.2. Machine learning, Deep Learning

ML, DL elhelyezése, 5. pdfem

2.2.1. Machine Learning

ML, ML in MIR

2.2.2. Deep Learning

DL, miben más mint az ML, architektúrák stb

3. fejezet

Módszertan

Lorem ipsum

3.1. Előfeldolgozás

adattisztítás, reprezentáció változtatás

3.2. Architektúra

3.3. Megvalósítás

3.3.1. Könyvtárak

4. fejezet

Adathalmaz

bevezető a datasetekről

4.1. Kiválasztási szempontok

pl. többszólam, ingyenesen elérhető, szerteágazó (not biased), stb

4.2. OpenMIC

4.3. MedleyDB

4.4. Egyebek

philharmonia orchestra irmas nsynth stb...

5. fejezet

Kísérletek, eredmények

Lorem ipsum

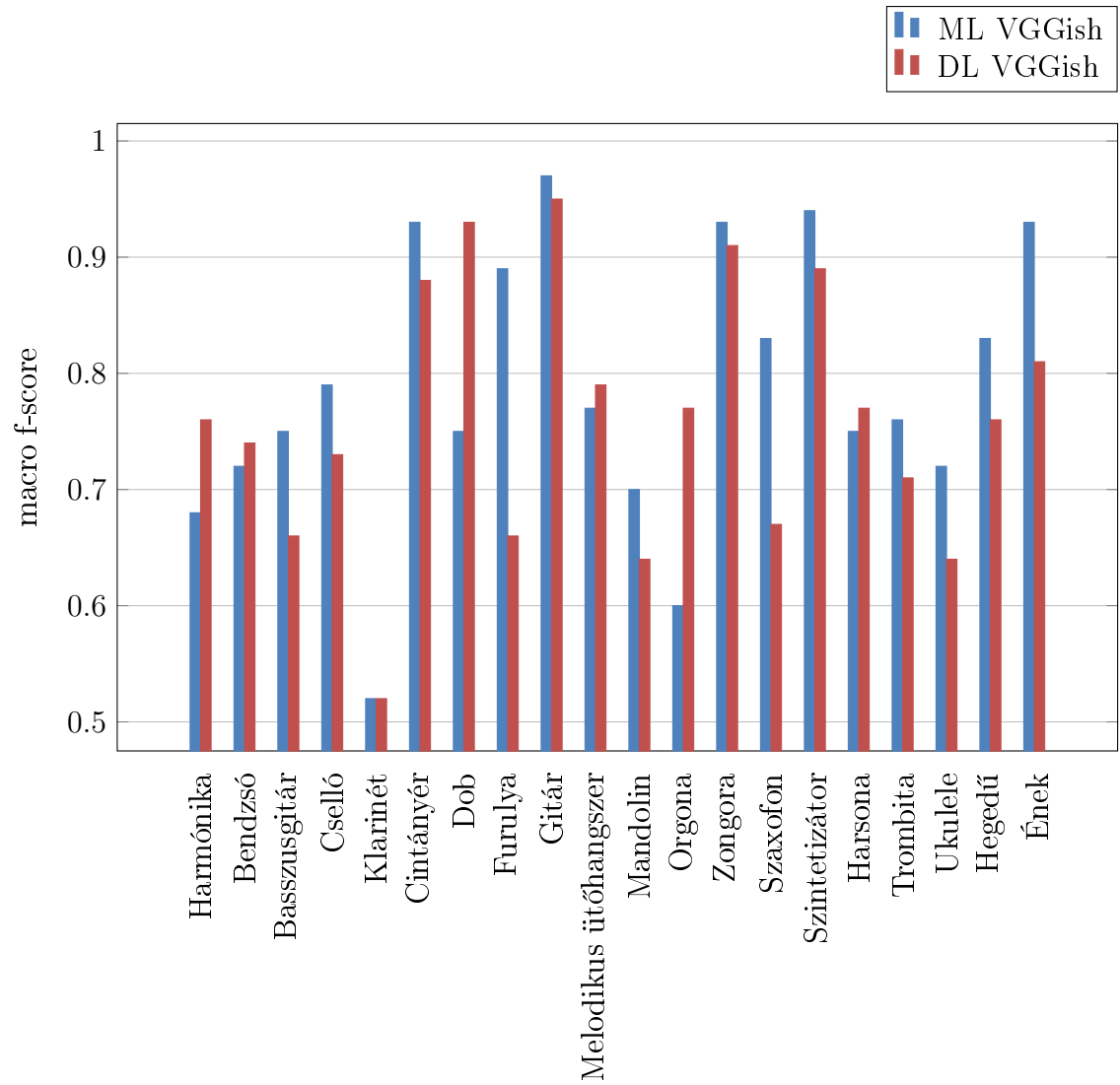
5.1. Mérőszámok

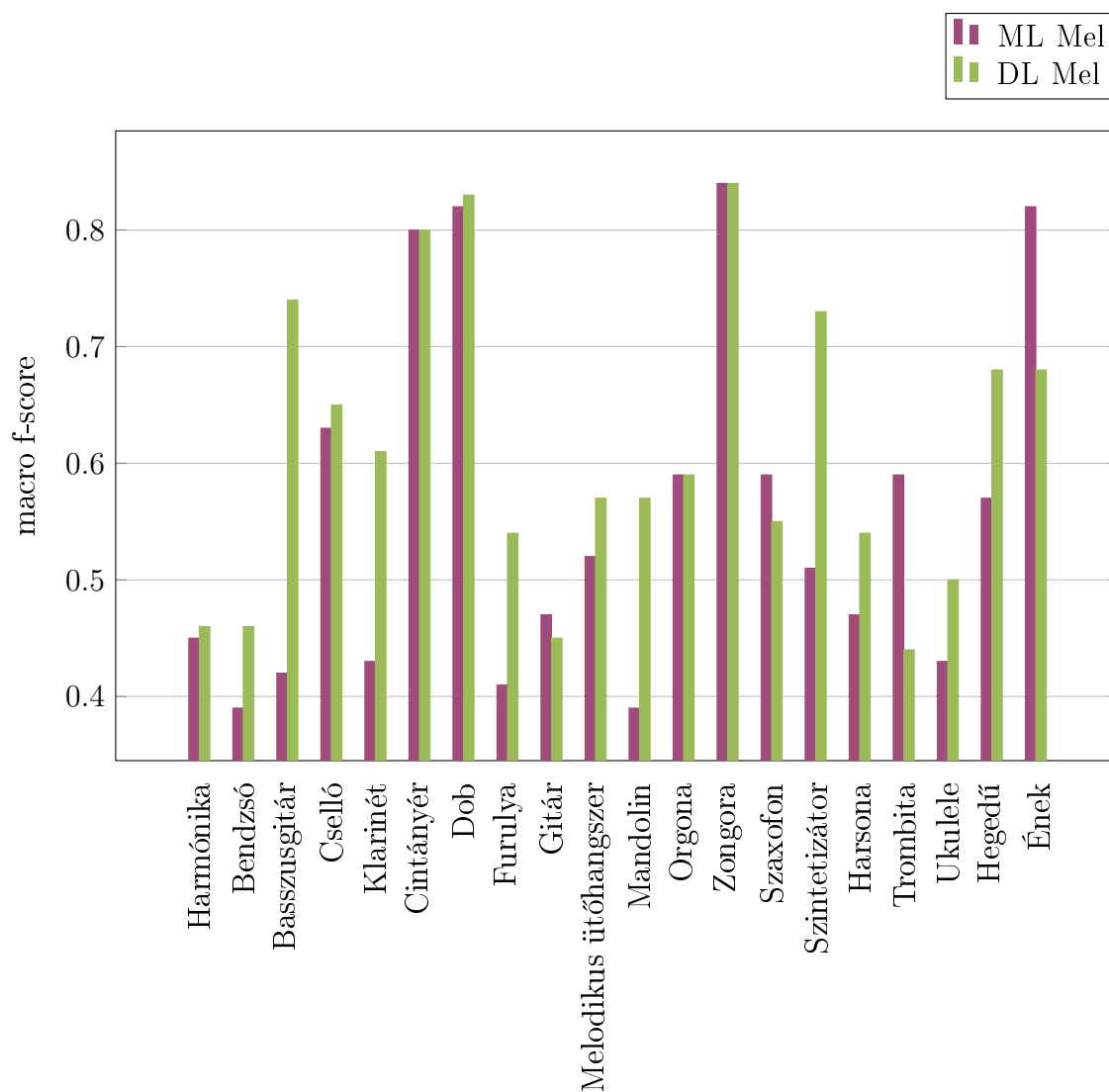
hogyan tudjuk megmérni a modell teljesítményét, pontosság, tanítás ideje, stb...

A következő metrikákat vizsgáltuk:

- asd
- asda

5.2. Eredmények





5.2.1. próba1

5.2.2. próba2

5.2.3. próba3

5.2.4. próba4

5.3. Összehasonlítás

5.3.1. saját próbálkozásaim

5.3.2. más munkákkal (hagyományos ML)

6. fejezet

Összegzés, kitekintés

Lorem ipsum dolor sit amet, consectetur adipiscing elit. In eu egestas mauris. Quisque nisl elit, varius in erat eu, dictum commodo lorem. Sed commodo libero et sem laoreet consectetur. Fusce ligula arcu, vestibulum et sodales vel, venenatis at velit. Aliquam erat volutpat. Proin condimentum accumsan velit id hendrerit. Cras egestas arcu quis felis placerat, ut sodales velit malesuada. Maecenas et turpis eu turpis placerat euismod. Maecenas a urna viverra, scelerisque nibh ut, malesuada ex.

Aliquam suscipit dignissim tempor. Praesent tortor libero, feugiat et tellus portitor, malesuada eleifend felis. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Nullam eleifend imperdiet lorem, sit amet imperdiet metus pellentesque vitae. Donec nec ligula urna. Aliquam bibendum tempor diam, sed lacinia eros dapibus id. Donec sed vehicula turpis. Aliquam hendrerit sed nulla vitae convallis. Etiam libero quam, pharetra ac est nec, sodales placerat augue. Praesent eu consequat purus.

Irodalomjegyzék

- [1] Peter Li, Jiyuan Qian és Tian Wang. “Automatic instrument recognition in polyphonic music using convolutional neural networks”. *arXiv preprint arXiv:1511.05520* (2015).
- [2] Yoonchang Han, Jaehun Kim és Kyogu Lee. “Deep convolutional neural networks for predominant instrument recognition in polyphonic music”. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 25.1 (2016), 208–221. old.