

Deep Learning a Vizuális Informatikában

Vizsga kérdések – 2023

Az alábbi kérdéssor a tárgy ismeret-ellenőrzéséhez kapcsolódó kérdésköröket ismerteti.

A vizsga e kérdéskörökre épül, ugyanakkor a kérdések nem feltétlen ebben a megfogalmazásban kerülnek a hallgatók elé.

A kérdések megalkotásánál fontos szempont volt, hogy részletesen szerepeljen, hogy mik azok a részletek, aminek az értését/tudását a vizsga során számonkérjük (ezért is meglehetősen hosszú a megfogalmazásuk).

Az államvizsga tárgyként e tárgyat választó hallgatók számára a kérdéssor a tárgyat 2023 tavaszi félévben végző hallgatók államvizsgatételeinek alapjaként is szolgál.

I. DEEP LEARNING ALAPJAI

1. Képfeldolgozási alapok

1. Alapfogalmak: számítógépes látás/grafika, képfeldolgozás. Alapvető feladatok: osztályozás, detektálás, szegmentálás (és fajtái), szintézis. Nehézségek.
2. Mesterséges intelligencia alapfogalma, fajtái. Képek felépítése és alapvető tulajdonságai.
3. Konvolúció művelete és alkalmazásai (szűrés, mintaillesztés). Konvolúció hatékony elvégzésének módjai (mátrixszorzás, szeparáció, frekvenciatartomány).

2. Neurális Hálók

4. Tanuló algoritmusok elve, felépítése, típusai. Gyakori hibák, over és underfitting. Legközelebbi szomszéd módszer működése és problémái.
5. Lineáris osztályozó (Perceptron modell) működése. Multi-Layer Perceptron felépítése, szigmoid nemlinearitás. Univerzális approximáció.
6. Hibafüggvények típusai, működése, előnyök/hátrányok. Mi a regularizáció és miért kell?
7. Optimalizáció: Gradiens módszer és kiegészítései. Newton módszer előnyei hátrányai.
8. Backpropagation: miért szükséges, hogyan működik. Neurális háló réteg általános interfésze.

3. Konvolúciós Hálók

9. Réteg típusok és azok tulajdonságai/paraméterei: Lineáris/FC, Konvolúciós, Pooling, Aktiváció/Nemlinearitás. Konvolúciós hálózatok általános felépítése.
10. Fontos architektúrák és azok mögött rejlő lényeges újítások/ötletek: AlexNet, VGG, Inception, ResNet, DenseNet.
11. Konvolúciós hálózatok vizualizációja, saliency, guided backpropagation. Adversarial példák generálása és azok elleni védekezés.

4. Deep Learning a Gyakorlatban

12. Konvergencia problémák okai és azok megoldása. Inicializációs módszerek, input normalizáció hatása.

13. Overfitting detektálása és elkerülése. Adatbázis felosztása és regularizációs módszerek. Batch Normalizáció, Dropout. Mi az adaugmentáció és a transzfer tanulás?

14. Hiperparaméterek meghatározása. Tanulási ráta gyors keresése és ütemezési stratégiák. Véletlen keresések módszerei: random vs grid vs Bayesi optimalizáció. NAS fogalma, ensemble modellek.

15. Hálók gyorsításának és tömörítésének módszerei (melyik módszer mit csinál): Pruning, Kvantálás, Weight Sharing. Hogyan lehet egy hálót már tanítás közben ritkítani? Szeparábilis konvolúció (térben és mélység mentén).

5. RNN és Transzformerek

16. Szekvenciális problémák fajtái. Hogyan vezethetünk dinamikát egy neurális háló rétegbe? RNN működése és tanítása (BPTT). Eltűnő gradiensek problémája.

17. LSTM réteg felépítése, működése. Hogyan oldja meg az eltűnő gradiensek problémáját? Puha figyelem elve és működése.

18. Self- és Cross-Attention réteg működése. Mi az a pozíció kódolás és miért szükséges? Előnyök/hátrányok az RNN-ekkel szemben.

19. Transformer encoder és decoder felépítése. Vision transzformerek elve, előnyei/hátrányai a konvolúciós architektúrákkal szemben. Swin transzformer felépítése és előnyei.

II. MAGASSZINTŰ FELADATOK

6. Detektálás, Követés

20. Lokalizáció módszere. Régió CNN és annak továbbfejlesztett változatai. Region Proposal network.

21. Single-shot detektorok. YOLO és változatai. Anchor box és non-maximum suppression fogalma. Spatial Pyramid pooling és Feature Pyramid Network elve és előnyei.

22. Anchor-free detektorok működése és előnyei. Gyakran használt mérőszámok, Recall, Precision, F-score, IoU, mAP.

23. Detektálás Transzformerek segítségével, DETR működése. Hogyan egészíthető ki ez a módszer követésre (TrackFormer, MOTR)?

7. Szegmentáció, Videóanalízis

24. Szegmentáció típusai: szemantikus, instance, panoptikus. FCN architektúrák felépítése és bennük használt trükkök, felskálázás típusai (unpooling, trConv, DUC). CRF elve és működése.

25. Instance és Panoptikus szegmentáció megoldására szolgáló architektúrák. Milyen egyéb feladatok oldhatók meg hasonló módon?

26. Videók feldolgozása: Integráció szintjei. 2D vs 3D konvolúció. Multimodális architektúrák. Optikai áramlás fogalma és becslése.

8. 3D feldolgozás

27. 3D információ előállítására alkalmas szenzorok és módszerek. SfM fogalma és elvégzése

neurális hálók segítségével. 3D LSTM működése. Monokuláris mélységbecslő architektúrák, Stereo transformer és NICE-SLAM.

28. 3D információ reprezentációja: Voxel rács, Pontfelhő, Mesh, kd-Fa, előnyök/hátrányok.

29. 3D tanulás módszerei: Voxel hálók, 2D vetületek, kd-Net, PointNet, PointCNN és MeshNet. Mi az a spatial transformer és mire jó?

III. NEM-FELÜGYELT TANULÁSI FORMÁK

9. Generatív hálók

30. Neurális stílus átültetés. Jellemző rekonstrukció, Gram-mátrix, textúra szintézis. A neurális stílus átültetés gyorsítási lehetőségei.

31. Generatív ellenséges háló (GAN) működésének alapötlete. GAN-ok tanítása. Túltanítás és a mode collapse jelensége.

32. Címkezett GAN-ok: Conditional GAN, InfoGAN. Wasserstein GAN bemutatása, Kullback-Leibler divergencia és a Wasserstein távolság összehasonlítása.

33. CycleGAN bemutatása: felépítés, tanítás (költségfüggvény), példák az alkalmazására.

10. Megerősítéses tanulás

34. Megerősítéses tanulás alapfogalma, Markov döntési folyamat és a benne foglalt fogalmak.

35. Érték és Q érték függvények, optimalitás feltétele, Bellman egyenlet és megoldása. Q tanulás és mély Q tanulás. Hátrányok, Experience replay.

36. Policy gradiens módszer elve és működése. Zaj redukciós módszerek: diszkontálás, baseline. Actor-critic és A2C módszerek. DDPG módszer alkalmazása.

37. Felfedezés problémája és stratégiák (e-greedy, entrópia). Kíváncsiság modell, world model architektúra. Kemény figyelem fogalma.

11. Ön-felügyelt tanulás

38. Ön-felügyelt tanulás fogalma és célja. Pre-text feladat fogalma, példák (képi és videó alapú). Videó színezés felhasználása követésre. Mi a hátránya ezeknek a módszereknek?

39. Generatív megoldások: Különböző Autoencoderes megoldások, Bidirectional GAN?

40. Kontrasztív tanulás fogalma, alapelve. Szimámi hálók felépítése. Különböző loss függvények: contrastive, Triplet, LSL, InfoNCE.

41. SimCLR framework működése, koszinusz hasonlóság. Tanítás módja. MoCo módszer elve és előnyei. CPC módszer működése. Előnyök/hátrányok a generatív módszerekkel szemben.

12. Neurális renderelés

42. Neurális radiancia mező (NeRF): milyen feladathoz fejlesztették ki, hogyan működik? Mi a különbség statikus és dinamikus szintér esetén?

43. Deformálható és kontrollálható szinterekhez alkalmazott neurális radiancia mezők. Hogyan lehet deformálható térre NeRF-et alkalmazni? Kontrollálható szintereknél milyen módon lehet kijelölni, hogy mely részleteket szeretnénk állíthatóvá tenni?

44. Neurális renderelés alkalmazásai: új nézetek szintézise, újvilágítás, arc és test újraalkotás.

13. Orvosi alkalmazások

45. Detekció és osztályozás orvosi képalkotásban. Hol alkalmazzák, milyen nehézségekkel jár,

milyen neurális háló architektúrákat lehet rájuk használni?

46. Szegmentáció orvosi képalkotásban. U-Net architektúra.

47. Kihívások a neurális hálók orvosi alkalmazásában: fejts ki 5 nehézséget, ami hátráltathatja a neurális hálók klinikai gyakorlatba helyezését!