

TASVIRNI TANISH UCHUN KLASSIFIKATSIYA MODELLARI, TENSORFLOW

Kenjayev Shahzodbek Yangiboy o'gli

Mirzo Ulug'bek nomidagi O'zbekiston Milliy Universiteti

shahzodkenjayevyangiboyevich@mail.com

Annotatsiya: Ushbu ishda konvolyutsion tarmoq modellarining keng ko'lamli tasvirni aniqlash metodlarining tensorflow tizimiga aniqligi ta'sirini o'rganildi. Ishning asosiy vazifasi - bu juda kichik (3×3) konvolyutsiya filtrlari bo'lgan arxitektura yordamida ortib borayotgan chuqurlikdagi tarmoqlarni to'liq baholash, bu chuqurlikni surish orqali avvalgi konfiguratsiyalarda sezilarli yaxshilanishga erishish mumkinligini ko'rsatadi.

Аннотация: В этой работе мы исследуем влияние глубины сверточной сети на ее точность в условиях крупномасштабного распознавания изображений с помощью Tensorflow. Наш основной вклад — это тщательная оценка сетей возрастающей глубины с использованием архитектуры с очень маленькими (3×3) сверточными фильтрами, которая показывает, что значительное улучшение по сравнению с конфигурациями предшествующего уровня техники может быть достигнуто за счет увеличения глубины до 16–19 весовых слоев. .

Abstract: In this work we investigate the effect of the convolutional network depth on its accuracy in the large-scale image recognition setting. Our main contribution is a thorough evaluation of networks of increasing depth using an architecture with very small (3×3) convolutionfilters, which shows that a significant improvement on the prior-art configurations can be achieved by pushing the depth to 16–19 weight layers.

Kalit so'zlar: Tensorflow, klassifikatsiya, model, inception, blok, ma'lumot, arxitektura, github, konvolyatsion tarmoq.

Keywords: Tensorflow, classification, model, inception, block, data, architecture, github, convolutional network.

KIRISH

Konvolyutsion tarmoqlar yaqinda katta hajmdagi tasvir va videoni aniqlashda katta muvaffaqiyatlarga erishdi (Krizhevskiy va boshq., 2012; Zeiler & Fergus, 2013; Sermanet va boshq., 2014; Simonyan & Zisserman, 2014) ImageNet (Deng va boshq., 2009) kabi katta ommaviy tasvir omborlari va GPU yoki keng miqyosli taqsimlangan klasterlar kabi yuqori unumdor kompyuter tizimlari (Dean va boshq., 2012) tufayli mumkin bo'ldi. Natijada, biz konvolyatsion tarmoq arxitekturasini yanada aniqroq ishlab chiqdik, ular nafaqat ILSVRC tasnifi va lokalizatsiya vazifalarida eng so'nggi

aniqlikka erishibgina qolmay, balki tasvirni aniqlash bo'yicha boshqa ma'lumotlar to'plamlarida ham qo'llanilishi mumkin. Nisbatan oddiy tizimlarning bir qismi sifatida foydalanilganda ham mukammal ishlashga erishiladi (masalan, aniq sozlashsiz chiziqli SVM tomonidan sinflangan chuqur xususiyatlar). Keyingi tadqiqotlarni osonlashtirish uchun ikkita eng yaxshi ishlaydigan model tanlab olindi. Ushbu maqolada konvulyatsion tarmoq konfiguratsiyalarini, tensorflow modeli yaratilishini va tasvirni klassifikatsiyalash bo'yicha o'qitish va baholash tafsilotlarini qisqacha yoritib beraman. To'liqlik uchun ilmiy ish mobaynida tensorflow ma'lumotlar bazasi ob'ektlarini mahalliyashtirish tizimimi yarataldi va baholandi va boshqa ma'lumotlar to'plamlari uchun juda chuqur xususiyatlarni umumlashtirishni muhokama qilindi.

Konvulyatsion tarmoq konfiguratsiyasi: konvulyatsion tarmoq chuqurligining oshishi natijasida yaxshilanishni adolatli sharoitda o'lchash uchun barcha Konvulyatsion tarmoq qatlam konfiguratsiyalarimiz Ciresan va boshqalar tomonidan ilhomlantirilgan bir xil tamoyillar asosida ishlab chiqilgan. (2011); Krizhevskiy va boshqalar. (2012). Ushbu bo'limda biz birinchi navbatda Konvulyatsion tarmoq- ning umumiy tartibini tasvirlaymiz konfiguratsiyalar (2. 1 - bo'lim) va keyin tafsilot the maxsus c _ konfiguratsiyalar _ _ _ baholash (2.2-bo'lim). Keyin dizayn tanlovlari muhokama qilinadi va Sektdagi oldingi tasvirlar bilan taqqoslanadi.

Arxitektura: konvulyatsion tarmoqga kirish 224×224 RGB o'lchamdagi o'rnatilgan rasmdir. Biz qiladigan yagona oldindan ishlov berish - bu o'quv to'plamida hisoblangan o'rtacha RGB qiymatini har bir pikseldan ayirish. Tasvir konvolyutsion (konv.) qatlamlar to'plamidan o'tkaziladi, bu erda biz juda kichik qabul qiluvchi maydonga ega filtrlardan foydalanamiz : 3×3 (bu chap/o'ng, yuqoriga/pastga, markaz). Konfiguratsiyalardan birida biz ham foydalanamiz 1×1 konvolyutsiya filtrlari, bu kirish kanallarining chiziqli o'zgarishi sifatida ko'rish mumkin (keyin chiziqli bo'lmagan). Konvolyutsiya qadami 1 pikselga o'rnatiladi; fazoviy to'ldirish konv.qatlam shunday kiritish kerakki, konvolyutsiyadan keyin fazoviy o'lchamlari saqlanib qoladi, ya'ni to'ldirish 3×3 konv. uchun 1 piksel. qatlamlar. Fazoviy birlashtirish beshta maksimal yig'ish qatlamlari tomonidan amalga oshiriladi, ular konv. qatlamlar. Maksimal birlashtirish 2×2 pikselli oynada, qadam 2 bilan amalga oshiriladi.

Konfiguratsiyalar: konvulyatsion tarmoq konfiguratsiyalari 1-jadvalda har bir ustunda bittadan keltirilgan. Quyida biz to'rlarga ularning nomlari bilan murojaat qilamiz (A–E). Barcha konfiguratsiyalar bo'limda keltirilgan umumiy dizaynga mos keladi. 2.1 va faqat chuqurlikda farqlanadi: A tarmog'idagi 11 ta og'irlik qatlamidan (8 konv. va 3 FC qatlamlari) E tarmog'idagi 19 ta vaznli qatlamlargacha (16 konv. va 3 FC qatlamlari).

ADABIYOTLAR TAHLILI VA METODOLOGIYA

Konvulyatsion tarmoq konfiguratsiyalari ILSVRC-2012 (Krizhevskiy va boshq., 2012) va ILSVRC-2013 musobaqalarining eng yaxshi natijalarga erishgan arizalarida qo'llanilganidan ancha farq qiladi (Zeyler va Fergus, 2013; Sermanet va boshq., 2014).. Birinchi konv.da nisbatan katta qabul qiluvchi maydonlardan foydalanish o'rniga. qatlamlar (masalan, 11×11 qadam 4 dyuym (Krizhevskiy va boshq., 2012) yoki 7×7 qadam 2 inç (Zeyler va Fergus, 2013; Sermanet va boshq., 2014)), biz juda kichik 3 dan foydalanamiz. \times Butun tarmoq bo'ylab 3 ta qabul qiluvchi maydonlar, ular har bir pikseldagi kirish bilan (1 qadam bilan) aylanadi. Ko'rinib turibdiki, ikkita 3×3 konv. qatlamli stek (orada fazoviy birlashmasdan) 5×5 samarali qabul qilish maydoniga ega; bunday qatlamlar 7×7 samarali qabul qiluvchi maydonga ega. Masalan, bitta 7×7 qatlam o'rniga uchta 3×3 konv.qatlamdan iborat stekdan foydalanib, biz nimaga erishdik ? Birinchidan, biz bitta qatlam o'rniga uchta chiziqli bo'lmagan to'g'rilash qatlamini o'z ichiga olamiz, bu esa qaror funktsiyasini diskriminativ qiladi. Hozirgi holatda 1×1 konvolyutsiya mohiyatan bir xil o'lchamli fazoga chiziqli proyeksiya bo'lsa ham (kirish va chiqish kanallari soni bir xil), to'g'rilash funktsiyasi tomonidan qo'shimcha chiziqli bo'lmaganlik kiritiladi. Shuni ta'kidlash kerakki, 1×1 konv. qatlamlar yaqinda Lin va boshqalarning "Tarmoqdagi tarmoq" arxitekturasida ishlatilgan. Kichik o'lchamli konvolyutsiya filtrlari ilgari Ciresan va boshqalar tomonidan ishlatilgan, lmodel ularning to'rlari biznikiga qaraganda sezilarli darajada chuqurroq va ular ILSVRC keng ko'lamli ma'lumotlar to'plamida baholanmagan. Goodfellow va boshqalar. Ko'cha raqamlarini aniqlash vazifasiga chuqur konvulyatsion tarmoqlar qo'llaniladi va chuqurlikning oshishi yaxshi ishlashga olib kelishini ko'rsatadi. GoogLeNet(Szegedy va boshq.,2014), ILSVRC- 2014 klassifikatsiyalash topshirig'ining eng yuqori samarali yozuvi bizning ilmiy ishimizdan mustaqil ravishda ishlab chiqilgan, biroq u juda chuqur konvulyatsion tarmoqlar va kichik konvolyutsiya filtrlaridan foydalaniladi. Biroq, ularning tarmoq topologiyasi biznikiga qaraganda murakkabroq va xususiyat xaritalarining fazoviy o'lchamlari birinchi qatlamlarda hisoblash hajmini kamaytirish uchun agressiv ravishda kamayadi. Bo'limda ko'rsatilgandek. 4.5, bizning modelimiz Szegedy va boshqalarnikidan ustundir.

Klassifikatsiya: Oldingi bo'limda biz tarmoq konfiguratsiyasining tafsilotlarini taqdim etdik. Ushbu bo'limda biz konvulyatsion tarmoq klassifikatsiyasini o'qitish va baholash tafsilotlarini tasvirlaymiz.

Tizimni o'qitish: Konvulyatsion tarmoq o'qitish jarayoni odatda Krizhevskiy va boshqalarga amal qiladi. Ya'ni, trening ko'p nomli logistik regressiya maqsadini impuls bilan mini-partiyali gradient tushishi yordamida optimallashtirish orqali amalga oshiriladi. Partiya hajmi 256 ga, momentum 0,9 ga o'rnatildi. Trening vaznning pasayishi va birinchi ikkita to'liq bog'langan qatlamlar uchun tartibni tark etish

(to'xtash darajasi 0,5 ga o'rnatildi) bilan tartibga solingan. 10^{-2} ga o'rnatildi va keyin tekshirish to'plamining aniqligi yaxshilanishni to'xtatganda 10 faktorga kamaydi. Hammasi bo'lib, o'rganish tezligi 3 marta kamaydi va 370 K iteratsiyadan (74 davr) keyin o'rganish to'xtatildi. Bizning taxminimizcha, (Krizhevskiy va boshq., 2012) bilan solishtirganda ko'proq parametrlar soniga va tarmoqlarimiz chuqurligiga qaramay, to'rlar (a) ko'proq chuqurlik va kichikroq konv. filtr o'lchamlari; (b) ma'lum qatlamlarni oldindan ishga tushirish. Tarmoq og'irliklarini ishga tushirish juda muhim, chunki noto'g'ri ishga tushirish chuqur tarmoqlardagi gradientning beqarorligi tufayli o'rganishni to'xtatib qo'yishi mumkin.

Tasvir o'lchami: S konvulyatsion tarmoq ma'lumotlari kesilgan izotropik o'lchamdagi o'quv tasvirining eng kichik tomoni bo'lishi (biz S ga o'quv shkalasi sifatida ham murojaat qilamiz). Kesish o'lchami 224×224 ga o'rnatilgan bo'lsada, printsiplial ravishda S 224 dan kam bo'lmagan istalgan qiymatni olishi mumkin: S = 224 uchun kesish o'quv tasvirining eng kichik tomonini to'liq qamrab olgan holda butun tasvir statistikasini oladi; $S \gg 224$ uchun kesish tasvirning kichik bir qismiga to'g'ri keladi, kichik ob'ekt yoki ob'ekt qismini o'z ichiga oladi. S o'quv shkalasini belgilash uchun ikkita yondashuvni ko'rib chiqamiz. Birinchisi, bir masshtabli o'qitishga to'g'ri keladigan x S ni tuzatish (namuna olingan modellar ichidagi rasm tarkibi hali ham ko'p masshtabli tasvir statistikasini ko'rsatishi mumkinligiga e'tibor bering). Tajribalarimizda biz ikkita qat'iy shkalada o'qitilgan modellari tahlil qilindi. Konvulyatsion tarmoq konfiguratsiyasini hisobga olgan holda, birinchi navbatda S = 256 yordamida tarmoqni tahlili o'tkazildi. S = 384 tarmog'ini o'qitishni tezlashtirish uchun u S = 256 bilan o'qitilgan og'irliklar bilan ishga tushirildi va biz kichikroq boshlang'ich o'rganish tezligi 10^{-3} dan foydalanildi. S ni o'rnatishning ikkinchi yondashuvi ko'p masshtabli trening bo'lib, unda har bir o'quv tasviri S ni ma'lum diapazondan tasodifiy tanlab olish yo'li bilan individual ravishda o'zgartiriladi [S,S]. Ob'ekt sin tasvirini skanerlash har xil o'lchamda bo'lganligi sababli, tajriba paytida buni hisobga olish kerak. Buni, shuningdek, masshtabni jitterlash orqali o'quv majmuasini kengaytirish sifatida ko'rish mumkin, bunda bitta model keng miqyosdagi ob'ektlarni tanib olishga o'rgatiladi. Tezlik uchun biz bir xil konfiguratsiyaga ega, bir xil konfiguratsiyaga ega bo'lgan barcha qatlamlarni aniq sozlash orqali ko'p masshtabli modellarni o'rgatish imkonini beradi.

Tizimni testlash: Sinov vaqtida o'qitilgan konvulyatsion tarmoq va kirish tasviri berilgan bo'lsa, u quyidagi tarzda klassifikatsiyalanadi. Birinchidan, u izotropik ravishda tasvirning oldindan belgilangan eng kichik tomoniga o'zgartiriladi, Q bilan belgilanadi (biz uni sinov shkalasi deb ham ataymiz). Shuni ta'kidlaymizki, Q o'quv shkalasi S ga teng bo'lishi shart emas. Keyin tarmoq o'zgartirilgan sinov tasviriga o'xshash tarzda zich ravishda qo'llaniladi. Natijada sinflar soniga teng kanallar soni va kirish tasvirining o'lchamiga bog'liq bo'lgan o'zgaruvchan fazoviy o'lchamlari

bo'lgan sinf reytinglari xaritasi. To'liq konvolyutsion tarmoq butun tasvir bo'ylab qo'llanilganligi sababli, sinov vaqtida bir nechta modellardan namuna olishning hojati yo'q (Krizhevskiy va boshq., 2012), bu unchalik samarali emas, chunki u har bir natija uchun tarmoqni qayta hisoblashni talab qiladi. Shuningdek, ko'p modellarni baholash turli konvolyutsiya chegarasi shartlari tufayli zich baholashni to'ldiradi : natijaga konvulyatsion tarmoq qo'llanganda, konvolved xususiyat xaritalari nol bilan to'ldiriladi, zich baholashda esa bir xil natija uchun to'ldirish tabiiy ravishda tasvirning qo'shni qismlaridan keladi. Konvolyutsiyalarga ham, fazoviy birlashtirishga ham, bu umumiy tarmoqning qabul qiluvchi maydonini sezilarli darajada oshiradi, shuning uchun ko'proq kontekst ushlanadi. Bizning fikrimizcha, amalda ko'p modellarni hisoblash vaqtining ko'payishi o'zini oqlamaydi potentsial daromad aniqlik, uchun ham baholash tarmoqlaridan foydalanish har bir masshtabda 50 ta model, 3 ta massivda jami 150 ta natija uchun, bu Szegedy va boshqalar tomonidan qo'llanilgan 4 ta o'lchovdagi 144 ta model bilan solishtirish mumkin.

NATIJARLAR

ILSVRC-2012 ma'lumotlar to'plamida tasvirlangan konvulyatsion tarmoq arxitekturalari tomonidan erishilgan tasvirlarni klassifikatsiyalash natijalarini taqdim etamiz (u ILSVRC 2012–2014 chal-lenges uchun ishlatilgan). Ma'lumotlar to'plami 1000 ta sinf tasvirlarini o'z ichiga oladi va uchta to'plamga bo'lingan: o'qitish (1.3 mln tasvir), tekshirish (50 K rasm) va test (ushlab qo'yilgan sinf yorliqlari bilan 100 K tasvir). Sinf ko'rsatkichlari ikkita o'lchov yordamida baholanadi. Birinchisi ko'p toifali sinflash xatosi, ya'ni noto'g'ri klassifikatsiyalangan tasvirlar nisbati ; ikkinchisi ILSVRCda qo'llaniladigan asosiy baholash mezonini bo'lib, tasvirlar nisbati sifatida hisoblanadi, shunday qilib asos-haqiqat toifasi bashorat qilingan eng yaxshi 5 toifadan tashqarida bo'ladi. Ko'pgina tajribalar uchun biz tekshirish to'plamidan test to'plami sifatida foydalandik. Sinov to'plamida ham ma'lum tajribalar o'tkazildi va ILSVRC-2014 tanloviga "VGG" jamoasi ishtirokchisi sifatida rasmiy ILSVRC serveriga topshirildi.

MUHOKAMA

Tensorflow to'plamidan olingan model bir qator muhim o'zgartirishlarni o'z ichiga oladi, bu bizga bitta tizimda o'rnatilgan bir nechta GPU-larda devalvatsiyani o'rganish imkonini beradi. Shuningdek, bitta blok to'liq o'lchamli tasvirlarni bir nechta masshtabda tekshirish murakkab hisoblanadi. Multi-GPU ma'lumotlar parallelligidan foydalanadi va o'quv tasvirlarining har bir partiyasini har bir GPUda parallel ravishda qayta ishlanadigan bir nechta GPU partiyalariga bo'lish orqali amalga oshiriladi. GPU partiyasi gradientlari hisoblangandan so'ng, to'liq partiyaning gradientini olish uchun ularning o'rtachasi olinadi. Gradient hisoblash GPUlar bo'ylab sinxronidir, shuning uchun natija bitta GPUda ishlash tezligi bir xil bo'ladi. Konvulyatsion tarmoq ta'limini

tezlashtirishning yanada murakkab usullari taklif qilingan bo'lsa-da (Krizhevskiy, 2014, ular tarmoqning turli qatlamlari uchun model va ma'lumotlar parallelligini qo'llaydilar, kontseptual jihatdan ancha sodda sxema allaqachon 3,75 baravar tezlashtirishni ta'minlaganligini aniqladik, bunda ma'lumotlar sinonimligi yahshi natija beradi. Bitta GPUdan foydalanish bilan solishtirganda, tayyor 4-GPU tizimi. To'rtta NVIDIA Titan Black GPU bilan jihozlangan tizimda bitta tarmoqni o'qitish arxitekturaga qarab 2-3 hafta davom etdi.

XULOSA

Alohida konvulyatsion tarmoq modellarining ishlashini tavsiflangan qatlam konfiguratsiyasi bilan bir miqyosda tahlil qilindi. Sinov tasvirining o'lchami quyidagicha o'rnatildi: $Q = S$ o'rnatilgan uchun S , va jitterli $S \in [S_{\min}, S_{\max}]$ uchun $Q = 0,5(S_{\min} + S_{\max})$. Birinchidan, shuni ta'kidlaymizki, mahalliy javobni normallashtirish (A-LRN tarmog'i) modelida normalizatsiya qatlamlarisiz yahshilanmaydi. Shunday qilib, chuqurroq arxitekturalarda normalizatsiyadan foydalanmaymiz (B-E). Klassifikatsiyalshash xatosi: konvulyatsion tarmoq chuqurligining ortishi bilan kamayishini kuzatamiz: A qatlamda 11 qatlamdan Eda 19 qatlamgacha. Shunisi e'tiborga loyiqki, bir xil chuqurlikka qaramay, C konfiguratsiyasi (u uchta 1×1 konv. o'z ichiga oladi). qatlamlar), 3×3 konv. ishlatadigan D konfiguratsiyasidan yomonroq ishlaydi. tarmoq bo'ylab qatlamlar. Bu shuni ko'rsatadiki, qo'shimcha chiziqli bo'lmaganlik yordam bersa ham (C B dan yaxshiroq), konv. yordamida fazoviy kontekstni olish ham muhim. ahamiyatsiz bo'lmagan retseptiv maydonlarga ega filtrlar (D C dan yaxshiroq). Bizning arxitekturamizdagi xatolik darajasi chuqurlik 19 qatlamga yetganda to'yingan bo'ladi, ammo undan ham chuqurroq modellar kattaroq ma'lumotlar to'plamlari uchun foydali bo'lishi mumkin. Bu ko'p masshtabli tasvir statistikasini olish uchun o'quv majmuasini masshtabni jitterlash orqali oshirish haqiqatdan ham foydali ekanligini tasdiqlaydi.

ADABIYOTLAR RO'YHATI

1. Bell, S., Upchurch, P., Snavely, N. va Bala, K. Yovvoyi tabiatda materiallarni kontekst bazasidagi materiallar bilan tanib olish. *CorR*, abs/1412.0623, 2014 yil.
2. Chat maydoni, K., Simonyan, K., Vedaldi, A. va Zisserman, A. Tafsilotlarda shaytonning qaytishi: konvolyutsion tarmoqlarga chuqur kirib borish. *Proc.da. _ BMVC.*, 2014 yil.
3. Cimpoi, M., Maji, S. va Vedaldi, A. To'qimalarni tanib olish va segmentatsiyalash uchun chuqur konvolyutsion filtr banklari. *CorR*, abs/1411.6836, 2014.
4. Ciresan, DC, Meier, U., Masci, J., Gambardella, LM va Schmidhuber, J. Tasvir tasnifi uchun moslashuvchan, yuqori samarali konvolyutsion neyron tarmoqlar. *IJCAIda*, 1237–1242-betlar, 2011 yil.
5. Dean, J., Korrado, G., Monga, R., Chen, K., Devin, M., Mao, M., Ranzato, M.,

- Senior, A., Taker, P., Yang, K., Le, QV va Ng, AY Katta miqyosda tarqalgan chuqur tarmoqlar. NIPSda, 1232–1240 - betlar, 2012 yil.
6. Deng, J., Dong, V., Socher, R., Li, L.-J., Li, K. va Fei-Fei, L. Imagenet : Keng miqyosli ierarxik tasvirlar bazasi. Proc.da. _ CVPR, 2009 yil.
 7. Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Chjan, N., Tzeng, E. va Darrell, T. Decaf: Umumiy vizual tanib olish uchun chuqur konvolyutsion faollashtirish xususiyati. *CorR*, abs/1310.1531, 2013 yil.
 8. Everingham, M., Eslami, SMA, Van Gool, L., Uilyams, C., Winn, J. va Zisserman, A. Paskalda vizual ob’ekt sinflari muammosi: retrospektiv. *IJCV*, 111(1):98–136, 2015.
 9. Fei-Fei, L., Fergus, R. va Perona, P. Bir nechta ta’lim misollaridan generativ vizual modellarni o’rganish: 101 ob’ekt toifalarida sinovdan o’tgan bosqichma-bosqich bayesian yondashuv. *IEEE CVPR generativ modelga asoslangan ko’rish ustaxonasida*, 2004 yil.
 10. Girshick, RB, Donahue, J., Darrell, T. va Malik, J. Rich ob’ektni aniq aniqlash va semantik segmentatsiya uchun ierarxialarni o’z ichiga oladi. *CorR*, abs/1311.2524v5, 2014. Proc.da nashr etilgan. CVPR, 2014 yil.