

4. Сухинин С.В. Гетерогенді ортада сигналдардың таралу ерекшеліктері.// Сб. ғылым. ең. Сем. Гомогенді және гетерогенді сұйықтықтар ағысының тұрақтылығы бойынша. Новосибирск, 1998, 98- 103 б.

5. Сухинин С.В. Кедергілердің периодтық тізбегінің толқынды, аномальды және шепчущая қасиеттері/ Индустриалды математика Сібір журналы, 1998, Т. 1, №2. 175 – 198 б.

ӘОЖ 004.93,004.89

АДАМНЫҢ ДИНАМИКАЛЫҚ ІС-ӘРЕКЕТІН ЖАДЫЛЫ RECURRENT NEURAL NETWORK АЛГОРИТМІ НЕГІЗІНДЕ ТАҢУ

Сансызбай Нұрболат
nurbolatsnk@gmail.com

Л.Н. Гумилев атындағы Еуразия ұлттық университеті, ақпараттық технологиялар факультетінің 2-курс магистранты
Ғылыми жетекшісі – Сатыбалдина Д.Ж.

Кіріспе

Жасанды сана(ЖС) адамның өмірін жеңілдетіп, оның қабілеттерін ұлғайта алады. Өз кезегінде ЖС шешімдер қабылдай алу үшін, алғы шарттарының бірі, шынайы өмірмен әрекеттесу интерфейстері мен құралдарының болуы, үлкен ақпараттар қорынан үлгілерді тани алуы болып табылады. Атап айтқанда, machine learning және deep learning әдістері кеңінен қолданысқа ие. Machine learning әдістері өз кезегінде бағдарламаның, бағдарламашысыз жазылу мүмкіндігін ие болса, Deep Learning әдістері сызықты және сызықты емес ақпараттан үлгілерді тануға мүмкіндік береді.

2020 жылғы жалпы әлемдік пандемия, ақпараттық технологияларда өңдеуші бағдарламаның болуы, онымен күресуге ықпал етуші шарттардың бірі екенің көрсетті. Соның ішінде орнатылған бейнекамераларда, адамның іс-әрекетін анализдеу, мақсатын анықтау маңызды. Мұнымен қоса, іс-әрекетті тану, денсаулықты бақылауда[1,2], ақылды үйлерде[3,4], адам-компьютер арақатынасында кеңінен қолданысқа ене алады.

Бейнелерден үлгілерді тануда, CNN(Convolutional neural network) орны ерекше. ImageNet деректер қорын жинаған Stanford университетінің оқытушысы Fei-Fei Li[5] еңбектерін атап айтуға болады. Алайда, CNN алгоритмі тек статикалық суреттерден керекті кілтті белгілерін шығару филтрлері арқылы, бір бағытта үлгілерді тани алады. Ал адамның іс-әрекеті динамикалық процесс болғандықтан, әрбір суреттегі адамның қалпы алдыңғы қалпын ескеру арқылы ғана болжам жасауға болады. Сондықтан адамның-іс әрекетін тануға арналған нейрондық желіге қойылатын талаптар келесідей:

1. Адамды бейнеден тану мүмкіндігінің болуы;
2. Танылған адамның орналасу қалпын анықтай алу;
3. Орналасу қалпы анықталған бірнеше суреттер тізбегінен адамның іс-әрекетін

тани алуы.

Адамның іс-әрекетін тану мақсатында, осы мақалада OpenPose, Recurrent Neural Network, LSTM алгоритмдерін қарастырамыз.

Алгоритмдер мен әдістерге шолу

Адамның іс-әрекетін тану негізгі екі қадамнан тұрады:

- Бейнеден адамның кілтті белгілерінің позициясын анықтау;
- Позициялар қатарынан іс-қимыл үлгісін тану.



Сурет-1. **Жоғарғы:** бірнеше адамның қалпын бағалау. Бір адамға тиесілі дене бөліктері аяқтың негізгі нүктелерін (үлкен саусақтар, кішкентай саусақтар мен өкшелер) қоса, өзара байланысты. **Төменгі сол жақ:** оң жақ шынтақ пен білезікті жалғайтын тиісті аяқ-қолдардың афинаждық өрісі (Part affinity fields – PAF’s). Түс бағдарды кодтайды. **Төменгі оң жақта:** әрбір пикселде екі өлшемді вектор әрбір PAF аяқ-қолдың орналасуы мен бағдарын кодтайды[6].

Бірінші қадамды шешуде OpenPose жүйесін қолдану шешілді. OpenPose - адам денесінің, қолдың, бет пен аяқтың (жалпы 135 негізгі нүкте) негізгі нүктелерін бөлек суреттерде ортақ табу үшін нақты уақыттағы алғашқы мультиадамдық ашық(open source) жүйе болып табылады. Windows, Mac, Ubuntu операциялық жүйелерін және CUDA (Nvidia GPU), OpenCL (AMD GPU) және тек-CPU (GPU-сіз) нұсқалары бар.

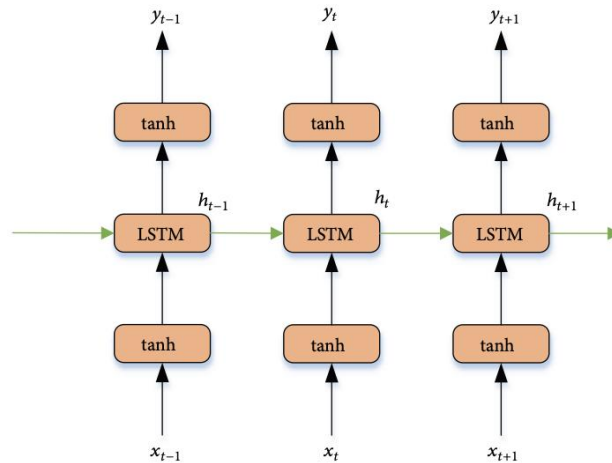
Екінші қадамды шешуде зерттеу келесідей бағыттарда жасалынды: 1D convolutional neural network(1D CNN), recurrent neural network(RNN) және long short-time memory(LSTM). Матрицалық көбейтулердің жеңілдігі және ұзындығы белгісіз тізбектер үшін үлгіні тану мүмкіндігінің болуы LSTM таңдауға себепші болды.

LSTM – бұл кеңейтілген RNN нейрондық желісі. Белгілі бір тереңдікте уақыт тізбектерінен мәселесін шеше алады. Сурет-2 көрсетілгендей, кірісінде $\{x_0, x_1, \dots, x_t, x_{t+1}, \dots\}$, шығысында, $\{y_0, y_1, \dots, y_t, y_{t+1}, \dots\}$ және ішкі деңгейі $\{h_0, h_1, \dots, h_t, h_{t+1}, \dots\}$ тізбектері арқылы анықталады. Ал, U, W, V сәйкесінше кіріс қабаттарының ішкі қабатқа, ішкі қабаттың ішкі қабатқа және ішкі қабаттың сыртқы қабатқа арақатынасын білдіретін салмақ матрицалары болып табылады. Жалпы жұмыс жасау принципі келесідей, кіріс қабаттың ішкі қабатқа және ішкі қабаттан ішкі қабатқа матрицалық түрлендіру арқылы және ішкі қабаттан сыртқы қабатқа активация(әдетте tanh, басқа да активация функцияларын қолдануға шектеу жоқ) функциясы арқылы мәнің аламыз. Формальды түрде ішкі қабат және сыртқы қабаттың мәндері келесідей анықталады:

$$h_i = \begin{cases} g(Ux_i + b_i^h) & i = 0 \\ g(Ux_i + Wh_{i-1} + b_i^h) & i = 1, 2, \dots \end{cases} \quad (1)$$

$$y_i = g(Vh_i + b_i^y) \quad i = 0, 1, \dots$$

бұл жердегі $g(\cdot)$ - tanh активация функциясы.



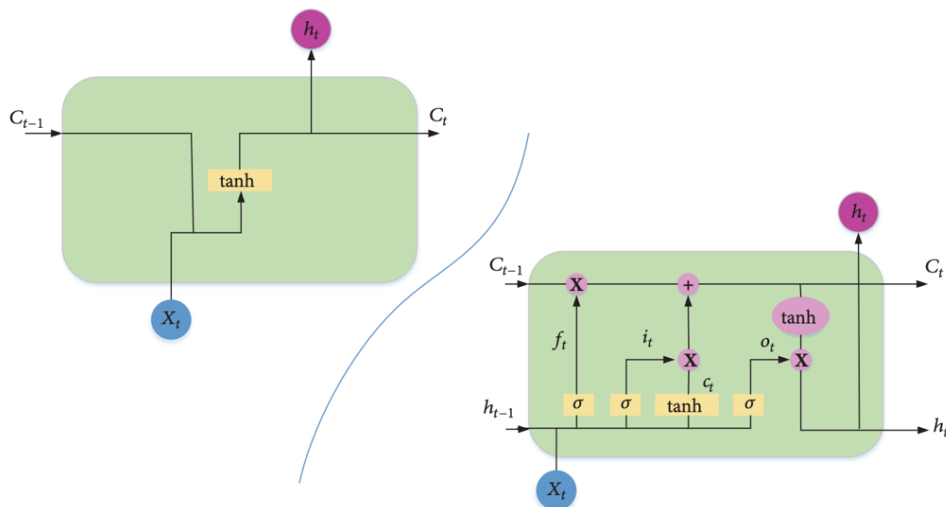
Сурет-2. Бір қабатты LSTM нейрондық желісінің құрылымы

RNN алдыңғы информацияға негізделе отырып, шығысты бағалай алады. Алайда, Ү. Bengio[7], RNN, “vanishing gradients” мәселесі әсерінен тек қысқа уақыт аралығын ғана жаттай алатынын анықтаған. Қарапайым конкатенация және активация функцияларынан тұратын RNN қарағанда, LSTM ұмыту қақпасы кіріс қақпасына қай ақпаратты жаттап, қай ақпаратты ұмытуын көрсетпегенше ұзақ мерзімді есте сақтау құрылымы бар.

LSTM кіріс, ұмыту және шығыс қақпалары, қай ақпараттың ұмытылуын, жатталуын, өзгеруін басқаруға арналған. Бұл керегінше ақпаратты селективті жібереді. Ақпараттың жіберілуін басқару үшін шығысындағы мәні [0,1] аралығында. LSTM ұяшығының әрбір параметрі t уақытында келесідей анықталады:

$$\begin{aligned}
 f_t &= \sigma(W_f[h_{t-1}, x_t] + b_f) \\
 i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i) \\
 \tilde{C}_t &= \tanh(W_c[h_{t-1}, x_t] + b_c) \\
 C_t &= f_t C_{t-1} + i_t \tilde{C}_t \\
 o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o) \\
 h_t &= o_t \tanh(C_t)
 \end{aligned}
 \tag{2}$$

бұл жердегі C_t – t уақытындағы ұяшық күйі.



Сурет-3. RNN және LSTM ұяшықтарының салыстырмасы. Үш σ операторлары(сол жақтан оң жаққа) сәйкесінше ұмыту(forget gate), кіріс(input gate), шығыс(output gate) қақпалары

Ақпаратты алдың ала өндеу

Мәліметтер Berkeley MHAD[8] деректер қорынан алынды. Беркли (MHAD) адамның мультимодальдық іс-әрекеттерінің деректер қоры бір қарт субъектіні қоспағанда, 7 ер адам мен 23-30 жас аралығындағы 5 әйел орындаған 11 іс-қимылды қамтиды. Барлық сыналушылар 660-қа жуық іс-әрекеттер тізбегін алып, әр әрекеттің 5 қайталауын орындалған, бұл жазудың жалпы уақытының шамамен 82 минутына сәйкес келеді. Сонымен қоса Kinect сенсорының мәліметтері де енгізілген. Мәліметтер қоры ғылыми мақсатта толықтай қолжетімді.

Біздің зерттеу тақырыбымызға келесідей әрекеттер алынды: секіру, қолмен секіру, бір қолын бұлғау, екі қолын бұлғау, жұдырықпен соққы беру, қол шапалақ.



Сурет-4. Berkeley MHAD-да қол жетімді барлық әрекеттердің жедел суреттері Kinect тереңдігінің деректерінен алынған нүктелердің тиісті бұлттарымен бірге көрсетіледі. Іс-әрекеттер (солдан оңға қарай): секіру, қол бұлғап секіру, бұғу, жұдырықпен соққы беру, екі қолмен бұлғау, бір қолмен бұлғау, алақанға құлату, лақтыру, отыру/тұру, отыру, тұру.

Жалпы бейнежазбаларға келесідей алдың ала өндеу жұмыстары жүргізілді:

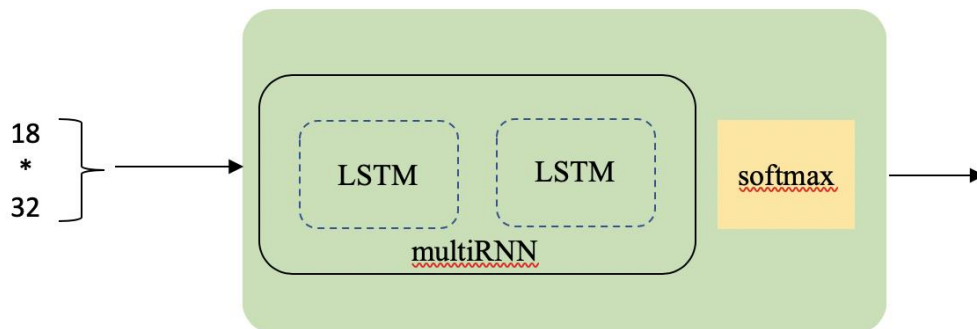
- Адамның іс әрекетіне байланысты кілтті белгілерінің x және y координатасын OpenPose жүйесі арқылы алу(18 joints);
- Алынған мәндерді json форматында сақтау;
- Шуы(мысалы, бір кадрда бірнеше адамның болу) бар кадрлардан тазарту, мәндерді нормализациялау және txt форматында сақтау.

Бейнежазбадағы алдың ала өндеуге жіберілген суреттер тізбегінің ұзындығы >200000 жоғары. Соның ішінде 5000-ға жуық суреттен шулар табылды. Жалпы бейнелер саны 1520.

Нейрондық желіні үйрету

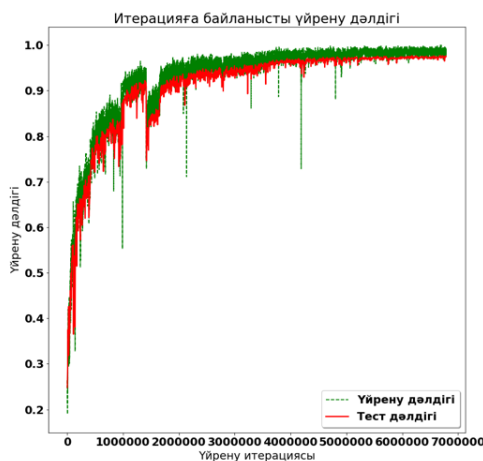
Нейрондық желіде екі LSTM тізбектей ұяшығы қолданылды. Оңтайландырушы ретінде Adam агоритмі және активация функциясы ретінде softmax таңдалынды. Әрбір бейне 32 кадрға бөлінді. Әрбір кадрға 18 кілтті белгілері жұбы белгіленді.

Нейрондық желіні үйрету процесі ~ 6787072 итерациядан тұрады. Тестілік ақпараттарды тануда нейрондық желі $\sim 97\%$ дәлдікті көрсетті. Сурет 6 көрсетілгендей, үйрету барысында градиентті түсудің шуының көптігін байқауға болады. Бұл үйретудің бір цикліне қатысатын ақпараттар ұзындығының кішілігін және үйрену көрсеткішінің аздығын көрсетеді.

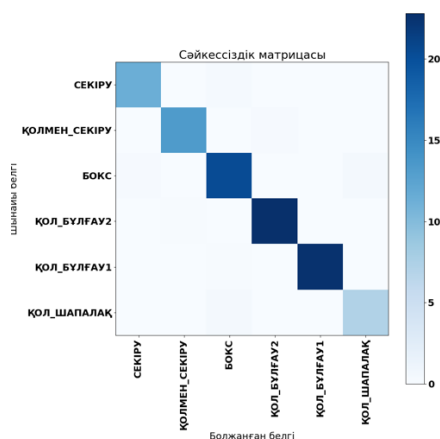


Сурет-5. Адамның іс-әрекетін тануға арналған LSTM негізді нейрондық желі құрылысы

Нейрондық желінің тану дәлдігі үлкен болғанымен үйрену уақыты баптауды талап етеді. Баптау барысында learning rate decay алгоритмі де қолданылып, үйрету көрсеткішінің әрбір цикл сайын кішірейуі ұйымдастырылды. Ақпараттар 80% үйретуге және 20% тестілеуге бөлінді.



Сурет-6. Итерацияға байланысты үйрену дәлдігінің өзгеру графигі



Сурет-6. Сәйкессіздік матрицасы

Қорытынды

Біз адамның іс-әрекетін тану бағытында LSTM алгоритмінің мүмкіндігі зерттелінді, нейрондық желі құрылды. Үйрену уақыты үлкен және градиентті түсуді есептеуде шудың көптігі болғанымен тестілік ақпараттар қорында нейрондық желі >90% дәлдікті көрсетті.

Алдағы жұмыстар үйрету уақытын қысқартуға, бейнебақылау камераларынан realtime адамның әрекетін тануды ұйымдастыруға бағытталады. Бейнекамералардағы адамның іс-әрекетін компьютердің түсіне алуы, адам мен компьютердің ара-қатынасын жаңа деңгейге алып келетіні анық.

Қолданылған әдебиеттер тізімі

1. H. Kalantarian, C. Sideris, B. Mortazavi, N. Alshurafa, and M. Sarrafzadeh, "Dynamic Computation Offloading for LowPower Wearable Health Monitoring Systems," IEEE Transactions on Biomedical Engineering, vol. 64, no. 3, pp. 621–628, 2017
2. A. Pantelopoulos and N. G. Bourbakis, "A survey on wearable sensor-based systems for health monitoring and prognosis," IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 40, no. 1, pp. 1–12, 2010.
3. S. H. Ahmed and D. Kim, "Named data networking-based smart home," ICT Express, vol. 2, no. 3, pp. 130–134, 2016.
4. S. Kumar, "Ubiquitous smart home system using android application," International Journal of Computer Networks & Communications, vol. 6, no. 1, pp. 33–43, 2014.
5. J. Niebles, B. Han, A. Ferencz, and L. Fei-Fei. Extracting moving people from internet videos. In ECCV, 2008.
6. Zhe Cao, Student Member, IEEE, Gines Hidalgo, Student Member, IEEE, Tomas Simon, Shih-En Wei, and Yaser Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields"
7. Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," IEEE Transactions on Neural Networks and Learning Systems, vol. 5, no. 2, pp. 157–166, 1994.
8. Ferda Ofli, Rizwan Chaudhry, Gregorij Kurillo, René Vidal, Ruzena Bajcsy, "Berkeley MHAD: A comprehensive Multimodal Human Action Database", 2013 IEEE Workshop on Applications of Computer Vision (WACV)
9. Schalk Wilhelm Pienaar, Reza Malekian, "Human Activity Recognition Using LSTM-RNN Deep Neural Network Architecture", arXiv:1905.00599
10. Zachary C. Lipton, John Berkowitz, Charles Elkan, "A Critical Review of Recurrent Neural Networks for Sequence Learning", arXiv:1506.00019
11. Vu Ngoc Thanh Sang, Nguyen Duc Thang, Vo Van Toi, Nguyen Duc Hoang, Truong Quang Dang Khoa, "Human Activity Recognition and Monitoring Using Smartphones", SpringerLink, 5th International Conference on Biomedical Engineering in Vietnam

УДК 621.391.812.3

ИССЛЕДОВАНИЕ МОДЕЛЕЙ КАНАЛОВ С МНОГОЛУЧЕВЫМ РАСПРОСТРАНЕНИЕМ СИГНАЛОВ

Сексембаева Манара Ануаровна

anuarkizi@gmail.com

докторант 1 курса ЕНУ им. Л.Н.Гумилева, Нур-Султан, Казахстан

Научный руководитель – Н. Н. Ташатов

Математической моделью канала называют совокупность системного оператора L и областей допустимых входных сигналов V_x и допустимых выходных сигналов V_y , [1, с.128].

Простейшей моделью многолучевого канала является описание интерференции двух лучей: первый луч распространяется по линии визирования, второй – интерферирует с первым после отражения от поверхности земли (рисунок 1).