# Թվային պատկերների իմաստաբանական վերլուծության մեթոդների մշակում և իրականացում

### ԹԵՄԱՅԻ ԱՐԴԻԱԿԱՆՈՒԹՅՈՒՆԸ

Թվային պատկերների իմաստաբանական վերուլուծությունը հանդիսանում է տեղեկատվական ճարտարագիտության (Information Engineering) հիմանական խնդրներից մեկը։ Այս խնդրի վերջնական լուծումը ենթադրում է, որ կստեղծվի համակարգ, որը կկարողանա ավտոմատ նկարագրել պատկերները այնպես ինչպես կարողանում է անել մարդը։ Այս խնդիրը առնչվում է այնպսիսի ոլորտներին ինչպիսիք են՝ մեքենայական ուսուցումը (Machine learning), վիճակագրոթյունը (Statistics), բնական լեզուների մշակումը (Natural language processing), համակարգչային տեսողությունը (Computer vision), ազդանշանների մշակումը (Signal Processing)` մասնավորապես պատկերնեի մշակումը (Image Processing), ինչպես նաև՝ ինֆորմացիայի տեսությունը (Information theory) և նեյրոկենսաբանությունը (Neurobiology)։

Թվային պատկերներ իմաստաբանական վելուծություն կատարելը ենթադրում է, որ համակարգի մուտքին տրվող պատկերը նկարագրող մատրիցի համար պետք է ձևավորվեն պատկերը նկարագրող այնպիսի մեծություններ, որոնք արտահայտված կլինինեն բնական լեզուների տարրերով՝ բառերով և նախասություններով։ Ընդորում պատկերը նկրագորղ առանձին վերցրած ֆիզիական մեծությունները չեն կարող համարվել պատկերի իմաստաբանական գնահատական՝ օրինակ որևէ պիքսելի գույնը կամ պատկերի պայծառությունը, բայց դրանք կարող են հանդիսանալ բառակապակցության, կամ նախադաության մաս՝ օրինակ *կարմիր* կատու, տղամարդը նստած է *պայծառ* սենյակում։ Խնդրի հիմանկան բարդությունը կայանում է նրանում, որ օրինակ պատկերների իմաստաբանական վերլուծության մաս կազմող պատկերում օբյետների հայտնաբերումը և տեղորոշումը չի ստացվում լուծել միայն պատկերնեի մշակման և կերպարների ճանաչման մեթոդներով (Pattern recognition): Խնդիրը կայանում է նրանում, որ պատկերում գտնվող օբյեկտները կախված իրար նկատմամբ դասավորվածությունից և մասշտաբից կարող են հանդիսանալ այլ առարկաներ, քան առանձին առաձին վերցված, օրիանկ առաձին վերցված կենդանու մազերը կարող են շատ նման լինինել սարդոստայինների։ Հենց այդ պատճառով էլ պատկերների մշակման սեգմենտացիայի ալգորիթմները դառնում են անզոր այս խնդրի լուծման համար։ Հենց նույն կերպ էլ հնարավոր չի լինում ճանաչել կերպարը, քանի որ անհայտ է, թե պատկերի, որ հատվածները պետք է ճանաչման համար համեմատվեն։ Այսպիսով ստացվում է, որ հնարավոր չէ կատարել օբյեկտնեի համեմատում, քանի որ հայտնի չեն պատկերում օբյտների եզրագծերը, իսկ եզրագծերի հայտնաբերման համար էլ անհրաժեշտ են օբյետների տեսակնաերը։ Այս ամենից եզրահանգում ենք այն, որ պատկերում օբյետների հայտնաբերում և տեղորոշում կատարող համակարգը պետք է օժտված լինի և ճանաչման և տեղորոշման հատկություններով՝ որոնք ուղղակիորեն կախված չլինեն իրարից։ Նմանատիպ խնդիրներ առաջանում են նաև պատկերը մեկ ընդանրական վերնագրով նկարագրելիս։ Վերջինումս առաջանում են նայև բնական լեզվի ճիշտ քերականությամբ նախադասության գենրացման հետ կապված խնդիրներ։

**Ատենախոսության նպատկան է** մշակել թվային պատկերների իմաստաբանական վերուլուծություն կատարող արագագործ ալգորիթմ (մոդել), որը կարող է նկարագրել պատկերը բնական լեզվի քերականությանը համապատասխանող նախադասությամբ՝ վերնագրով ինչպես նաև ճանաչել և տեղորոշել պատկերում գտնվող օբյեկտները։ Մշակված ալգորիթմի համար ստեղծել այնպիսի ծրագրային հմակարգեր, որոնք հնարավոր լինեն օգտագործել համակրգչի կենտրոնական պրոցեսորի վրա, ինչպես նաև անհրաժեշտության դեպքում օգտագործել գրաֆիկական քարտի պրոցեսորը հաշվարկների արագացման համար։

**Հետազոտության մեթոդներ**․ Թվային պատկերների իմաստաբանական վերլուծության մեթոդների մշակման համար օգտոգրծվել են պատկերները դասկարգող, օբյետներ հայտնաբերող և տեղորոշող խորը փաթույթային նեյրոնային ցանցեր, պատկերի վերնագիրը բնական լեզվով արտահյտող ռեկուրենտ նեյրոնային ցանցեր։ Ստեղծվել են ծրագրային համակարգեր մոդելների ուսուցման, գնահատման և թեստավորման համար։ Ծրագրային համակարգերում օգտագործվել են զուգահեռ հաշվարկների մեթոդներ և գրադարաններ։ Ծրագրային համակարգերը նախագծված են այնպես, որպեսզի օգտագործեն և կենտրոնական պրոցեսորի բոլոր միջուկների և համակարգչում տեղադրված գրաֆիկական քարտի հաշվողական հզորությունները։ Ուսուցման և գնահատման համար օգտագործվել են բաց պատկերների բազաներ և համակարգի ճշտության ստուգման հայտնի ալգորիթմներ։

**Արդյունքների գիտական նորությունները**

* Մշակվել է իմաստաբանական նման պատկերներների փնտրման ալգորիթմ հիմնված խորը փաթույթային նեյոնային ցանցի բարձր մակարդակի հատկությունների վրա։
* Մշակվել է պատկերի LSTM ռեկուրենտ նեյրոնային ցանցի մոդիֆիկացված տարբերակ, որը ցուցաբերել է ավելի բարձր ճշտություն պատկերների վերնագրի գեներացման խնդրում։
* Մշակվել է պատկերի վերնագիրը գեներացնող ռեկուրենտ ցանց հիմնվելով օբյեկտներ հայտաբերող և տեղորոշող խորը փաթույթային նեյրոնային ցանցի բարձր մակարդկի հատկությունների վրա։
* Մշակվել են մեծ արագագործություն ունեցող ծրագրային համակարգեր, որոնք կարողանում են օգտագրծել համակարգչի ողջ հաշվողական հզորությունները՝ կենտորնական և գրաֆիկական քարտի պրոցեսորների բոլոր միջուկները, այդպիսով հնարավոր դարձնելով մեծ նեյրոնային ցանցերի ուսուցումը, գնահատումը և թեսթավորումը։
* Մշակվել է մեծ արագագործություն ունեցող ծրագրային համակարգ, որը կատարում է պատկերի վերնագրի գեներացում, օբյեկտների ճանաչում և տեղորոշում։

**Աշխատանքի կիրառական նորությունները։** Մշակված պատկերի իմաստաբանական վերլուծություն կատարող համարգը կարելի է ներդնել փնտրող համակարգերում նկարագրությամբ կամ բանալի բառերով պատկերների որոնման համար։ Տեսողության խնդիրներ ունեցող մարդկանց կաորղ է օգնել վիրտուալ, ինչպես նաև իրական աշխարհում, տալով այդտեղ գտնվող առարակաների, իրադարձությունների մասին ինֆորմացիա։ Կարելի է ավտոմատացնել որոշ հսկողության համակարգեր ներդնելով մշակված ծրագրային համակարգերը։ Լայն կիրռություն կարող է ունենալ տարբեր արտադրական և ուրիշ նշանակության համար նախատեսված ռոբոտներում։

* 1. *N. Hayrapetyan, R. Hakobyan, A. Poghosyan, V. Gabrielyan, “Border Surveillance Using UAVs with Thermal Camera”, NATO Science for Peace and Security Series – D: Information and Communication Security, vol. 47, pp. 219 – 225,* *2016.*
	2. *A. Poghosyan, H. Sarukhanyan, "Image Visual Similarity Based on High Level Features of Convolutional Neural Networks", Mathematical Problems of Computer Science, vol. 45, pp. 138-142, 2016.*
	3. *A. Poghosyan, H. Sarukhanyan, “Long Short-Term Memory with Read-only Unit in Neural Image Caption Generator”, 11-th International Conference Computer Science and Information Technologies, (CSIT-11), 2017, pp. 368-371.*
	4. *A. Poghosyan, H. Sarukhanyan, “Short-term memory with read-only unit in neural image caption generator”, 11-th International Conference Computer Science and Information Technologies, Revised Selected Papers, IEEE Xplore,* [*10.1109/CSITechnol.2017.8312163*](https://doi.org/10.1109/CSITechnol.2017.8312163)*, Electronic ISBN: 978-1-5386-2830-0, Print on Demand(PoD) ISBN: 978-1-5386-2831-7, pp. 162- 167, 2017*
	5. *A. Poghosyan, H. Sarukhanyan, “RNN with Additional Constant Memory for Image Caption Generation Task”, International Academy Journal Web of Scholar, 4(13), Vol.1, pp. 3–7, 2017.*
	6. *A. Poghosyan, “Image Caption Generation and Object Detection via a Single Model”, Mathematical Problems of Computer Science vol. 48, pp. 42–49, 2017.*
	7. *A. Poghosyan, H. Sarukhanyan, "* *Image Caption Generation model based on Object Detector", Mathematical Problems of Computer Science, (in progress).*