

ABSTRACT

**of the dissertation for the degree of Doctor of Philosophy (PhD) in specialty
“6D070400 - Computing Systems and Software”**

AITIMOV ASKHAT KABIDOLLAYEVICH

Analysis and development of algorithm and method for object recognition

Relevance of the work: In the modern era, the volume of visual data produced daily across various domains—ranging from security surveillance to healthcare imaging and human–computer interaction—has grown exponentially. This proliferation of images and videos creates an urgent need for intelligent systems capable of accurately identifying and interpreting visual information. Object recognition, a fundamental component of computer vision, underpins many such technologies by enabling automated perception and decision-making similar to the human visual system. Reliable object recognition is critical for developing applications such as facial authentication, gesture-based interfaces, and intelligent monitoring systems that require precise, real-time analysis of visual scenes.

Despite significant progress achieved through deep learning, modern recognition systems remain heavily dependent on large, diverse, and well-annotated datasets. In practice, however, obtaining extensive labeled data is often infeasible due to privacy concerns, annotation costs, and limited sample diversity—particularly in specialized domains such as biometrics and gesture recognition. Models trained under these constraints typically suffer from overfitting, poor generalization, and instability when deployed in real-world environments. Furthermore, the reliance of contemporary architectures on high-performance computational hardware limits their accessibility for embedded or low-resource applications.

These challenges highlight the scientific and practical importance of developing algorithms that can operate efficiently with limited data and modest computational resources. Research in this area supports the broader goal of sustainable artificial intelligence by promoting models that are interpretable, lightweight, and adaptable to constrained environments. Creating such methods not only advances the theoretical understanding of machine perception but also facilitates their implementation in real-life systems that must perform reliably despite restricted datasets and computational limitations.

Aim of the research work: To develop an algorithm and method for face and hand gesture recognition efficiently capable of achieving high accuracy under limited dataset conditions.

Objectives of the research work:

1. To employ limited datasets frequently utilized in research endeavors.
2. To develop a lightweight and efficient recognition approach for face and hand gesture identification that ensures high accuracy and computational efficiency in data-constrained environments.

Research methods: The research was conducted using a combination of analytical and experimental methods. The analytical component involved the examination of existing approaches to object, face, and hand gesture recognition under limited data conditions, as well as the identification of factors affecting recognition accuracy and generalization. The experimental component focused on developing and evaluating a lightweight recognition algorithm based on PCA, Modified Triplet Similarity Loss, and Projection. For the experimental evaluation, two limited datasets were employed: the ORL face dataset and the HandReader gesture dataset, both commonly used in academic research for small-scale recognition studies. PCA was applied to reduce the dimensionality of input features and eliminate redundant information, while the Modified Triplet Similarity Loss function was used to improve feature discrimination and clustering in the embedding space. Finally, a projection-based classification stage was implemented to enhance recognition stability under limited-data conditions.

Scientific Novelty: The scientific novelty of this research lies in the development of a lightweight object recognition algorithm and method that integrate PCA, a Modified Triplet Similarity Loss, and a Projection-based classification stage to ensure accurate performance under limited dataset conditions for face and hand gesture recognition.

Scientific, Practical, and Theoretical Significance of the Study: The theoretical significance of this research lies in the development of a hybrid methodological framework that integrates PCA, a Modified Triplet Similarity Loss, and Projection-based classification into a unified recognition system. This combination expands the theoretical understanding of how dimensionality reduction, metric learning, and projection principles can be jointly optimized to improve recognition accuracy under limited dataset conditions. The study contributes to the theoretical foundations of machine learning by demonstrating how loss function modification and feature-space projection can enhance the discriminative capability of embeddings in data-scarce environments. Moreover, it enriches the methodological tools available to researchers by proposing a lightweight, interpretable recognition architecture that balances performance, computational efficiency, and generalization, an aspect often overlooked in existing deep learning frameworks. The practical significance of the work is reflected in its direct applicability to face and hand gesture recognition tasks where data availability is constrained. The proposed algorithm can be effectively implemented in systems requiring reliable

recognition with minimal computational resources, such as embedded platforms, human–computer interaction interfaces, and security applications. Its use of the ORL and HandReader datasets demonstrates adaptability to small-scale, real-world scenarios, confirming its potential for integration into low-resource recognition systems. The method’s lightweight structure ensures ease of deployment and accessibility, offering practical value for organizations and researchers operating under data or hardware limitations. Overall, the research supports the broader goal of developing efficient and sustainable recognition technologies suitable for diverse and data-constrained environments.

Scientific Hypotheses, Main Findings, and Significance of Results: The central hypothesis of the dissertation is that high-accuracy object recognition under limited data conditions can be achieved by integrating dimensionality reduction, adaptive metric learning, and projection-based feature regularization within a single lightweight framework. This hypothesis was confirmed both analytically and experimentally using the ORL and HandReader datasets, demonstrating that accurate and stable recognition does not depend solely on large-scale data or deep architectures but rather on the optimized interaction of feature extraction and similarity learning components.

The key findings submitted for defense include:

1. The development of the PCA–TP (Principal Component Analysis – Triplet Projection) algorithm and method, which combine linear dimensionality reduction with a modified metric-learning approach for efficient and interpretable object recognition.
2. The introduction of a Modified Triplet Similarity Algorithm, which integrates cosine similarity, an adaptive angular margin, and projection consistency regularization. This unified loss formulation enhances the discriminative structure of embeddings, balances intra-class compactness with inter-class separation, and ensures stable convergence during training.
3. The experimental validation of the proposed method and algorithm, showing recognition accuracies of 90–100% on the ORL and HandReader datasets, while maintaining low computational complexity and a reduced number of parameters.

The scientific novelty of the research lies in the creation of a unified and lightweight recognition algorithm and method that merge PCA-based dimensionality reduction with a modified triplet similarity algorithm—integrating cosine similarity, adaptive margin adjustment, and projection consistency into a single optimization framework. This approach provides a new methodological perspective on metric learning for data-scarce environments by combining geometric regularization and linear feature projection.

The significance of the results is reflected in both theoretical and practical contributions. Theoretically, the research advances understanding of how angular similarity and projection-based regularization enhance learning stability and discrimination under limited data. Practically, the proposed algorithm and method can be effectively implemented in resource-constrained or embedded systems for applications such as biometric authentication, gesture-based control, and intelligent monitoring.

Publications and Author's Contribution: The results of this dissertation were published in a series of scientific papers that collectively reflect the progressive development of the author's research on object, face, and gesture recognition under limited-data conditions. Each study contributed to the formulation of the proposed PCA–TP algorithm and the overall methodological framework of this dissertation.

The paper “**Sign Language Dactyl Recognition Based on Machine Learning Algorithms**” (Eastern-European Journal of Enterprise Technologies, 2021, Vol. 4, Issue 2, pp. 58–72) presented a system for recognizing Kazakh dactylic sign language using machine learning methods. In this research, several algorithms—Random Forest, Support Vector Machine, and Extreme Gradient Boosting—were applied to a dataset of 5,000 images per gesture, achieving classification accuracies of 98.86%, 98.68%, and 98.54%, respectively. The study demonstrated the potential of classical machine learning techniques for high-accuracy sign language recognition. The author contributed to algorithmic development, coding, and evaluation of classifiers, and later adapted the methodological approach of this study to the experimental framework of the dissertation.

The article “**2D Face Recognition Using PCA and Triplet Similarity Embedding**” (Bulletin of Electrical Engineering and Informatics, 2023, Vol. 12, Issue 1, pp. 580–586) introduced the PCA–TP algorithm that integrates dimensionality reduction and triplet-based metric learning with a projection classifier. The author developed the entire conceptual design, preprocessing, and experimentation independently. The algorithm achieved up to 98.75% recognition accuracy on the ORL dataset with an 80/20 train-test split and demonstrated stable performance across lower ratios. This study forms the main experimental foundation of the dissertation, serving as the basis for the final PCA–TP model combining PCA, Modified Triplet Similarity Loss, and Projection-based classification.

The work “**PCA and Projection-Based Hand Gesture Recognition**” (Journal of Theoretical and Applied Information Technology, 2021, Vol. 99, Issue 15, pp. 3801–3811) focused on vision-based gesture recognition using the HandReader dataset. The author proposed, implemented, and experimentally validated a method combining PCA for feature reduction, triplet embedding for discriminative feature learning, and projection-based classification. Despite allocating only 20% of the

data for training and 80% for testing, the algorithm achieved up to 95% accuracy, significantly outperforming traditional methods such as KNN and MSRC. The entire research concept, algorithmic development, and analysis were conducted by the author, and its methodology was later extended to the dissertation's PCA-TP framework.

The conference paper “**An Enhanced Face Recognition Method for Lighting Problem**” (15th International Conference on Electronics, Computer and Computation, ICECCO 2019, pp. 1–4) investigated preprocessing techniques to improve face recognition accuracy under illumination variations. The combination of Histogram Equalization and Adaptive Single Scale Retinex (ASSR) increased the recognition rate from 10% (without preprocessing) to 90% when both techniques were applied. The author contributed to the implementation of PCA and experimental evaluation. The results of this work were used in the dissertation's preprocessing methodology for face recognition tasks.

The earlier study “**Gesture Recognition Based on Sparse Reconstruction**” (14th International Conference on Electronics, Computer and Computation, ICECCO 2018, pp. 206–212) examined the use of the MSRC algorithm for gesture recognition and compared it with KNN and Random Forest classifiers. The experiments—fully designed and implemented by the author—showed that MSRC achieved superior recognition rates (up to 97.5%) and faster execution time on the HandReader dataset. The results from this study directly influenced the dissertation's focus on combining sparse representation and projection-based learning for small datasets.

The publication “**Comparative Analysis of Recognition Algorithms for Hand Gestures on the Basis of Various Representations of Images**” (Herald of the Kazakh-British Technical University, 2019, Vol. 16, Issue 1, pp. 50–54) provided an extensive comparison of KNN, Logistic Regression, and Decision Tree classifiers using raw pixel and histogram-based features. The experiments, carried out entirely by the author, showed that the Logit classifier with raw pixel input achieved 98% recognition accuracy but required higher computation time. This work supported the dissertation's conclusion that feature representation significantly affects performance under limited-data conditions and motivated the later use of PCA for compact feature encoding.

In the study “**Information Technology for Sign Language Alphabet Units Modeling and Recognition**” (Bulletin of KazNITU, 2020, Vol. 5, pp. 364–371), a cross-platform gesture modeling and recognition technology was developed using virtual hand models and standardized protocols such as YAML and PostgreSQL for gesture databases. The system supported multi-language gesture sets and demonstrated the feasibility of cross-platform recognition architectures. The author contributed to the CNN-based recognition component and system design.

The article “**Comparative Analysis of Face Recognition Algorithms in the Problem of Visual Identification**” (Suleyman Demirel University Bulletin:

Natural and Technical Sciences, 2021, Vol. 54, Issue 1, pp. 61–67) analyzed classical algorithms such as EigenFace and Gabor filters for visual identification systems. The author supervised the project, guiding master's students in developing the research methodology, data preparation, and implementation. Although this study was not directly used in the dissertation experiments, its comparative results helped contextualize classical approaches in relation to the PCA–TP algorithm.

Finally, the work “**Methods for Handling Frames of Hand Gesture Recognition**” (Suleyman Demirel University Bulletin: Natural and Technical Sciences, 2021, Vol. 54, Issue 1, pp. 45–51) explored preprocessing methods for static gesture images extracted from video streams, including smoothing, binarization, rotation, and background modification based on the skin color model. The author supervised and reviewed the experimental process, ensuring methodological rigor. The techniques presented in this paper informed the preprocessing procedures applied to the datasets used in the dissertation's experiments.

Volume and Structure of the Dissertation: The dissertation consists of 94 pages and is organized into several chapters that comprehensively cover both the theoretical foundation and the experimental validation of the research. It begins with an Introduction, followed by an extensive Literature Review exploring the history, relevance, and challenges of pattern recognition, with a specific focus on limited datasets and modern techniques such as data augmentation, GANs, transfer learning, and few-shot learning. The Research Methodology chapter details the datasets used (including the HandReader and ORL datasets), the algorithms applied (PCA, modified triplet loss, projection), and the design of the proposed PCA–TP algorithm. The Experiment and Results chapter presents a series of hand gesture and facial recognition experiments, including data preprocessing, classifier evaluations, and comparative analysis of various methods.

The dissertation includes 156 bibliographic references, establishing a strong scientific and academic basis for the work. To enhance clarity and provide visual insight into complex processes, 31 diagrams and 18 tables are presented throughout the text. This structured approach ensures a logical and comprehensive presentation of the research, which focuses on developing an effective object recognition algorithm under limited data conditions, with the novel contribution being the PCA–TP method.

АННОТАЦИЯ

«6D070400 – Есептеу техникасы және бағдарламалық қамтамасыз ету» мамандығы бойынша философия докторы (PhD) ғылыми дәрежесін алу үшін диссертацияға

АЙТИМОВ АСХАТ КАБИДОЛЛАЕВИЧ

Объектілерді тану алгоритмі мен әдісін талдау және әзірлеу

Жұмыстың өзектілігі: Қазіргі таңда әртүрлі салаларда – қауіпсіздік жүйелерінен бастап медицинадағы бейне диагностика мен адам–компьютер өзара әрекеттесуіне дейін – визуалды деректер көлемі экспоненциалды түрде артып келеді. Күн сайын алынатын суреттер мен бейнелердің көбеюі визуалды ақпаратты дәл және жедел талдай алатын интеллектуалды жүйелерді қажет етеді. Компьютерлік көрудің негізгі құрамдас бөлігі болып табылатын объектіні тану технологиясы машиналарға адам көзіне ұқсас түрде қоршаған ортаны тануға және шешім қабылдауға мүмкіндік береді.

Бүгінгі таңда объектілерді тану технологиясы бетті идентификациялау, қол қимылдарын тану және бейнебақылау сияқты нақты қолданбалы салаларда маңызды рөл атқарады.

Алайда қазіргі терең оқыту әдістері үлкен, әртүрлі және мұқият таңбаланған деректер жиынтығына тәуелді. Мұндай деректерді жинау көп жағдайда жеке мәліметтердің құпиялылығына, сараптамалық аннотацияның күрделілігіне және деректердің шектеулігіне байланысты мүмкін болмай жатады. Нәтижесінде модельдер шағын деректермен оқытылғанда, артық үйрену (overfitting) мен жалпылаудың төмендігі байқалады, ал жоғары өнімді есептеу ресурстарына тәуелділік оларды кіріктірілген немесе төмен қуатты жүйелерде қолдануды қиындатады.

Сондықтан шектеулі деректер жағдайында тиімді жұмыс істейтін, түсініктемелі және есептеу жағынан жеңіл алгоритмдерді жасау – қазіргі жасанды интеллекттің маңызды ғылыми бағыты болып табылады. Мұндай зерттеулер теориялық тұрғыдан машиналық көрудің дамуына, ал практикалық тұрғыдан – шектеулі ресурстар жағдайында сенімді жұмыс істейтін тану жүйелерін құруға ықпал етеді.

Зерттеу жұмысының мақсаты: Шектеулі деректер жағдайында жоғары дәлдікпен жұмыс істейтін бет және қол қимылдарын тануға арналған тиімді алгоритм мен әдісті әзірлеу.

Зерттеу мақсаттары:

1. Ғылыми зерттеулерде жиі қолданылатын шағын деректер жиынтықтарын пайдалану.
2. Деректердің шектеулі көлемінде жоғары дәлдік пен есептеу тиімділігін қамтамасыз ететін жеңіл тану әдісін әзірлеу.

Зерттеу әдістері: Зерттеу жұмысы аналитикалық және эксперименттік әдістердің үйлесімі арқылы орындалды. Аналитикалық бөлік объектілерді, бет пен қол қимылдарын танудың қазіргі тәсілдерін талдау және олардың шектеулі деректер жағдайындағы тиімділігін бағалауға бағытталды. Эксперименттік бөлік PCA (негізгі компоненттерді талдау), модификацияланған Triplet Similarity Loss және проекцияға негізделген классификация әдістерін біріктіретін жаңа жеңіл алгоритмді әзірлеу мен бағалауға арналды. Зерттеу барысында ORL бет-әлпет деректер жиынтығы және HandReader қол қимылдары жиынтығы қолданылды. PCA өлшемділікті азайту және ақпараттық қайталануды жою үшін пайдаланылды; модификацияланған triplet loss функциясы белгілердің дискриминативтілігін арттыруға бағытталды; соңғы кезеңде проекцияға негізделген классификатор тану тұрақтылығын қамтамасыз етті.

Зерттеудің ғылыми жаңалығы: Зерттеу нәтижесінде PCA, модификацияланған Triplet Similarity Loss және проекцияға негізделген классификацияны біріктіретін, шағын деректер жағдайында жоғары дәлдікті қамтамасыз ететін жеңіл объектіні тану алгоритмі мен әдісі әзірленді.

Зерттеудің ғылыми, практикалық және теориялық маңыздылығы: Теориялық тұрғыдан бұл жұмыс өлшемділікті азайту, метрикалық оқыту және проекция қағидағтарын біріктіретін жаңа гибридті тану тәсілін ұсынады. Мұндай интеграция шектеулі деректер жағдайында танудың дәлдігін арттыруға мүмкіндік береді және метрикалық оқыту теориясын дамытады. Практикалық тұрғыдан ұсынылған әдіс аз есептеу ресурстарын қажет ететін, деректер көлемі шектеулі қолданбалы жүйелерде – кіріктірілген құрылғыларда, қауіпсіздік және адам–компьютер интерфейстерінде – тиімді қолданыла алады. Зерттеу нәтижелері нақты шағын деректер жиынтықтарында тексеріліп, әдістің жоғары дәлдігі мен есептеу тиімділігі дәлелденді.

Ғылыми болжам, негізгі нәтижелер және олардың маңызы: Жұмыстың негізгі гипотезасы – жоғары дәлдіктегі тануды шектеулі деректер жағдайында өлшемділікті азайту, адаптивті метрикалық оқыту және проекциялық реттеу қағидаларын біріктіретін біртұтас құрылым арқылы жүзеге асыруға болады. Бұл болжам теориялық және тәжірибелік тұрғыдан дәлелденіп, ұсынылған PCA–TP алгоритмінің ORL және HandReader жиынтықтарында 90–100% дәлдікке қол жеткізетіні көрсетілді.

Қорғауға ұсынылатын негізгі нәтижелер:

1. PCA–TP алгоритмі мен әдісі әзірленді – ол өлшемділікті азайту мен модификацияланған метрикалық оқытуды біріктіріп, тиімді және түсініктемелі тануды қамтамасыз етеді.
2. Косинустық ұқсастық, адаптивті бұрыштық шек және проекциялық тұрақтылықты біріктіретін модификацияланған Triplet Similarity алгоритмі енгізілді.
3. Ұсынылған әдіс тәжірибелік түрде дәлелденді – ORL және HandReader жиынтықтарында 90–100% дәлдікке жетіп, есептеу күрделілігі төмен екені анықталды.

Зерттеудің ғылыми жаңалығы PCA негізіндегі өлшемділікті азайту әдісін модификацияланған triplet similarity алгоритмімен біріктіретін, бірыңғай әрі жеңіл тану алгоритмі мен әдістемесін әзірлеуде жатыр. Бұл әдісте косинустық ұқсастық, адаптивті шекті мәнді реттеу және проекциялық тұрақтылық бір оңтайландыру құрылымына біріктірілген. Мұндай тәсіл аз деректер жағдайындағы метрикалық оқытуға жаңа әдіснамалық көзқарас береді, себебі ол геометриялық реттеу мен сызықтық белгілерді проекциялауды біріктіреді. Зерттеу нәтижелерінің маңыздылығы теориялық және практикалық тұрғыдан көрініс табады. Теориялық тұрғыда жұмыс шектеулі деректер жағдайында бұрыштық ұқсастық пен проекциялық реттеу оқу процесінің тұрақтылығын және белгілердің ажыратушылық қабілетін арттыратынын дәлелдейді. Практикалық тұрғыдан ұсынылған алгоритм мен әдіс биометриялық аутентификация, қол қимылдарын тану және интеллектуалды мониторинг сияқты қолданбаларда ресурсы шектеулі немесе кіріктірілген жүйелерге тиімді түрде енгізілуі мүмкін.

Жарияланымдар және автордың үлесі: Бұл диссертацияның нәтижелері шектеулі деректер жағдайында нысандарды, бет бейнелерін және қол қимылдарын тану бойынша зерттеулердің кезең-кезеңімен дамуын көрсететін бірқатар ғылыми мақалаларда жарияланды. Әрбір жарияланым ұсынылған PCA–TP алгоритмі мен осы диссертацияның жалпы әдістемелік негізін қалыптастыруға өз үлесін қосты.

“Sign Language Dactyl Recognition Based on Machine Learning Algorithms” (Eastern-European Journal of Enterprise Technologies, 2021, Vol. 4, Issue 2, pp. 58–72) атты мақалада машинамен оқыту әдістерін пайдалану арқылы қазақ дактильді ым тілі белгілерін тану жүйесі ұсынылды. Бұл зерттеуде Random Forest, Support Vector Machine, және Extreme Gradient Boosting алгоритмдері әр қимылға 5000 кескіннен тұратын деректер жиынтығына қолданылды. Нәтижесінде сәйкесінше 98.86%, 98.68%, және 98.54% дәлдікке қол жеткізілді. Бұл зерттеу жоғары дәлдікті ым тілін тануда классикалық машиналық оқыту әдістерінің әлеуетін көрсетті. Автор алгоритмдерді әзірлеу,

бағдарламалау және бағалау процестеріне үлес қосты, сондай-ақ осы жұмыстың әдістемелік тәсілдерін диссертацияның тәжірибелік бөлігіне бейімдеді.

“2D Face Recognition Using PCA and Triplet Similarity Embedding” (Bulletin of Electrical Engineering and Informatics, 2023, Vol. 12, Issue 1, pp. 580–586) мақаласында өлшемділікті азайту мен triplet негізіндегі метрикалық оқытуды біріктіретін және проекциялық классификатормен толықтырылған PCA–TP алгоритмі ұсынылды. Зерттеу идеясы, алдын ала өңдеу және тәжірибелер толығымен автордың тарапынан орындалды. Алгоритм ORL деректер жиынтығында 80/20 үйрету/тестілеу қатынасында 98.75% дәлдікке жетті және төменгі қатынастарда да тұрақтылық көрсетті. Бұл зерттеу диссертацияның негізгі тәжірибелік базасын құрап, PCA, Modified Triplet Similarity Loss, және Projection-based classification біріктірілген соңғы PCA–TP моделін қалыптастыруға негіз болды.

“PCA and Projection-Based Hand Gesture Recognition” (Journal of Theoretical and Applied Information Technology, 2021, Vol. 99, Issue 15, pp. 3801–3811) еңбегі HandReader деректер жиынтығын пайдаланып, көру негізіндегі қол қимылдарын тануға арналды. Автор PCA арқылы белгілерді азайту, triplet embedding арқылы дискриминативті белгілерді үйрену және projection негізіндегі классификацияны біріктіретін әдісті ұсынды, жүзеге асырды және тәжірибелік түрде дәлелдеді. Деректердің тек 20%-ы үйретуге, ал 80%-ы тестілеуге бөлінсе де, алгоритм 95% дәлдікке жетіп, KNN және MSRC әдістерінен асып түсті. Зерттеу идеясы мен талдау жұмыстары толықтай автор тарапынан жүргізіліп, бұл әдіс кейін диссертациядағы PCA–TP шеңберіне енгізілді.

“An Enhanced Face Recognition Method for Lighting Problem” (15th International Conference on Electronics, Computer and Computation, ICECCO 2019, pp. 1–4) конференциялық мақаласында жарықтандыру жағдайларының әсерін азайту арқылы бет бейнелерін тану дәлдігін арттыру әдістері зерттелді. Histogram Equalization және Adaptive Single Scale Retinex (ASSR) әдістерін біріктіру нәтижесінде дәлдік 10%-дан (алдын ала өңдеусіз) 90%-ға дейін артты. Автор PCA енгізу және тәжірибелік сынақтарды орындауға үлес қосты. Бұл жұмыстың нәтижелері диссертациядағы бет бейнелерін алдын ала өңдеу әдісіне енгізілді.

“Gesture Recognition Based on Sparse Reconstruction” (14th International Conference on Electronics, Computer and Computation, ICECCO 2018, pp. 206–212) атты ертеректегі зерттеуде қол қимылдарын тану үшін MSRC алгоритмін пайдалану және оны KNN мен Random Forest классификаторларымен салыстыру қарастырылды. Барлық тәжірибелер мен алгоритмдік іске асыру толықтай автор тарапынан орындалды. Алынған нәтижелер HandReader деректер жиынтығында 97.5% дәлдік пен жылдам орындалу уақытын көрсетті. Бұл жұмыс диссертациядағы аз деректер жағдайында sparse representation мен projection-based learning тәсілдерін біріктіру бағытына тікелей ықпал етті.

“Comparative Analysis of Recognition Algorithms for Hand Gestures on the Basis of Various Representations of Images” (Herald of the Kazakh-British Technical University, 2019, Vol. 16, Issue 1, pp. 50–54) мақаласында KNN, Logistic Regression, және Decision Tree алгоритмдері raw pixel және histogram белгілері негізінде салыстырылды. Барлық тәжірибелер мен талдауларды автор жүргізді. Logit классификаторы raw pixel енгізулерімен 98% дәлдікке жетіп, бірақ есептеу уақыты ұзақ болғаны анықталды. Бұл зерттеу деректерді ұсыну тәсілінің тиімділігіне әсерін дәлелдеп, кейінгі PCA негізіндегі белгілерді ықшамдау шешіміне түрткі болды.

“Information Technology for Sign Language Alphabet Units Modeling and Recognition” (Bulletin of KazNITU, 2020, Vol. 5, pp. 364–371) зерттеуінде YAML және PostgreSQL стандартталған протоколдары арқылы виртуалды қол моделіне негізделген кросс-платформалық ым моделдеу және тану технологиясы әзірленді. Жүйе көптілді ым жиынтықтарын қолдады және кросс-платформалық архитектуралардың тиімділігін көрсетті. Автор CNN негізіндегі тану компонентін және жүйелік жобалауды әзірлеуге қатысты.

“Comparative Analysis of Face Recognition Algorithms in the Problem of Visual Identification” (Suleyman Demirel University Bulletin: Natural and Technical Sciences, 2021, Vol. 54, Issue 1, pp. 61–67) мақаласында EigenFace және Gabor филтрлері сияқты классикалық алгоритмдер визуалды идентификация жүйелерінде талданды. Автор бұл жобаны жетекшілік етіп, магистранттарға зерттеу әдістемесін қалыптастыру, деректерді дайындау және тәжірибелік іске асыру бойынша нұсқаулық берді. Бұл жұмыс диссертацияның тәжірибелік бөлігіне тікелей енгізілмегенімен, классикалық әдістер мен PCA–TP алгоритмінің салыстырмалы контексін қалыптастыруға мүмкіндік берді.

Соңында, **“Methods for Handling Frames of Hand Gesture Recognition”** (Suleyman Demirel University Bulletin: Natural and Technical Sciences, 2021, Vol. 54, Issue 1, pp. 45–51) мақаласында бейне ағындардан алынған статикалық қол қимылдарының кескіндерін өңдеу әдістері – тегістеу, бинаризация, айналдыру және фонды тері түсі моделіне негіздеп түзету – қарастырылды. Автор тәжірибелік процестерді ғылыми тұрғыдан қадағалап, нәтижелердің әдістемелік дәлдігін қамтамасыз етті. Бұл жұмыста сипатталған алдын ала өңдеу әдістері диссертациядағы деректер жиынтықтарын дайындау процесінде қолданылды.

Осы тоғыз жарияланым біртұтас ғылыми бағыттың дәйекті дамуын көрсетіп, соңында ұсынылған PCA–TP алгоритмі мен оның әдістемелік негізін қалыптастырды. Автордың үлесі ғылыми зерттеу циклының барлық кезеңдерін — идеяны қалыптастыру, алгоритмді әзірлеу, тәжірибелік тексеру және нәтижелерді жариялау — қамтып, шектеулі деректер жағдайында тиімді, түсінікті және ықшам тану жүйесін құруға әкелді.

Диссертацияның көлемі мен құрылымы: Диссертация 94 беттен тұрады және зерттеудің теориялық негіздері мен тәжірибелік нәтижелерін жан-жақты қамтитын бірнеше бөлімдерден құралған. Жұмыс Кіріспеден басталып, әрі

карай әдеби шолу бөлімінде үлгілерді тану саласының тарихы, өзектілігі және негізгі мәселелері талданады. Бұл бөлімде шектеулі деректер жағдайындағы зерттеу әдістеріне, соның ішінде data augmentation, GANs, transfer learning, және few-shot learning сияқты заманауи тәсілдерге ерекше назар аударылған. Зерттеу әдістемесі бөлімінде қолданылған деректер жиынтықтары (HandReader және ORL), пайдаланылған алгоритмдер (PCA, modified triplet loss, projection) және ұсынылған PCA–TP алгоритмінің құрылымы егжей-тегжейлі сипатталған. Эксперимент және нәтижелер бөлімінде қол қимылдарын және бет бейнелерін тануға арналған тәжірибелер, деректерді алдын ала өңдеу, классификаторларды бағалау және әртүрлі әдістердің салыстырмалы талдауы келтірілген. Диссертацияда 156 библиографиялық дереккөз пайдаланылған, бұл жұмыстың ғылыми және академиялық негізінің мықтылығын дәлелдейді. Сонымен қатар, күрделі процестерді көрнекі түрде бейнелеу үшін мәтін бойына 31 сызба және 18 кесте енгізілген. Мұндай құрылымдық тәсіл зерттеу материалын қисынды және толық түрде ұсынуға мүмкіндік береді. Диссертацияның негізгі мақсаты — шектеулі деректер жағдайында тиімді жұмыс істейтін нысандарды тану алгоритмін әзірлеу, ал оның негізгі ғылыми үлесі PCA–TP әдісімен ұсынылған.

АННОТАЦИЯ

**диссертации на соискание ученой степени доктора философии (PhD) по
специальности
"6D070400 - Вычислительная Техника и Программное
Обеспечение"**

АЙТИМОВ АСХАТ КАБИДОЛЛАЕВИЧ

Анализ и разработка алгоритма и метода распознавания объектов

Актуальность исследования: В современную эпоху объем визуальных данных, ежедневно создаваемых в различных областях от систем видеонаблюдения до медицинской визуализации и взаимодействия человека с компьютером стремительно возрастает. Этот поток изображений и видео формирует острую необходимость в интеллектуальных системах, способных точно распознавать и интерпретировать визуальную информацию. Распознавание объектов, являясь фундаментальной составляющей компьютерного зрения, лежит в основе множества технологий, обеспечивающих автоматическое восприятие и принятие решений, аналогичных человеческому зрительному восприятию. Надежное распознавание объектов играет ключевую роль в разработке таких приложений, как системы аутентификации по лицу, интерфейсы на основе жестов и интеллектуальные системы мониторинга, требующие точного анализа визуальных сцен в реальном времени.

Несмотря на значительный прогресс, достигнутый благодаря методам глубокого обучения, современные системы распознавания по-прежнему сильно зависят от больших, разнообразных и тщательно аннотированных наборов данных. На практике получение таких обширных размеченных данных часто затруднено из-за требований конфиденциальности, высокой стоимости разметки и ограниченного разнообразия выборок особенно в специализированных областях, таких как биометрия и распознавание жестов. Модели, обученные в условиях дефицита данных, как правило, страдают от переобучения, низкой обобщающей способности и нестабильности при применении в реальных условиях. Кроме того, зависимость современных архитектур от высокопроизводительного вычислительного оборудования ограничивает их применение в встраиваемых и ресурсно-ограниченных системах.

Эти проблемы подчеркивают научную и практическую значимость разработки алгоритмов, способных эффективно функционировать при ограниченных данных и умеренных вычислительных ресурсах. Исследования

в данном направлении способствуют реализации более широкой цели устойчивого развития искусственного интеллекта, предполагающего создание интерпретируемых, лёгких и адаптируемых моделей. Разработка таких методов не только углубляет теоретическое понимание машинного восприятия, но и обеспечивает возможность их практической реализации в системах, которые должны работать надежно даже при ограниченных объемах данных и вычислительных мощностях.

Цель исследовательской работы: Разработать алгоритм и метод распознавания лиц и жестов рук, способный эффективно обеспечивать высокую точность в условиях ограниченных наборов данных.

Задачи исследовательской работы:

1. Использовать ограниченные наборы данных, наиболее часто применяемые в научных исследованиях.
2. Разработать легкий и эффективный подход к распознаванию лиц и жестов рук, обеспечивающий высокую точность и вычислительную эффективность в условиях ограниченных данных.

Методы исследования: Исследование выполнено с использованием сочетания аналитических и экспериментальных методов.

Аналитическая часть включала обзор существующих подходов к распознаванию объектов, лиц и жестов рук в условиях ограниченного объема данных, а также определение факторов, влияющих на точность и обобщающую способность моделей распознавания. Экспериментальная часть была направлена на разработку и оценку легковесного алгоритма распознавания, основанного на PCA, Modified Triplet Similarity Loss и Projection. Для экспериментальной оценки использовались два ограниченных набора данных: ORL (для распознавания лиц) и HandReader (для распознавания жестов рук), которые широко применяются в академических исследованиях маломасштабных задач распознавания. Метод PCA применялся для снижения размерности входных признаков и устранения избыточной информации, в то время как функция Modified Triplet Similarity Loss использовалась для повышения дискриминативности признаков и улучшения их кластеризации в пространстве эмбедингов. На заключительном этапе была реализована классификация на основе проекции, что позволило повысить устойчивость распознавания в условиях ограниченных данных.

Научная новизна: Научная новизна исследования заключается в разработке легковесного алгоритма и метода распознавания объектов, объединяющего

PCA, Modified Triplet Similarity Loss и Projection-based classification в единую систему, обеспечивающую высокую точность работы при ограниченных наборах данных в задачах распознавания лиц и жестов рук.

Научная, практическая и теоретическая значимость исследования:

Теоретическая значимость работы заключается в создании гибридной методологической структуры, которая объединяет PCA, Modified Triplet Similarity Loss и Projection-based classification в единую систему распознавания. Такое сочетание расширяет теоретическое понимание совместной оптимизации методов снижения размерности, метрического обучения и проекционного подхода для повышения точности распознавания при ограниченных данных. Исследование вносит вклад в теоретические основы машинного обучения, демонстрируя, как модификация функции потерь и проекция в пространстве признаков усиливают дискриминативные свойства эмбедингов в условиях недостатка данных. Кроме того, работа обогащает методологический инструментарий исследователей, предлагая легковесную и интерпретируемую архитектуру распознавания, сочетающую производительность, вычислительную эффективность и способность к обобщению — аспект, часто игнорируемый в современных глубинных моделях. Практическая значимость исследования проявляется в его применимости к задачам распознавания лиц и жестов рук при ограниченном количестве данных. Предложенный алгоритм может быть эффективно реализован в системах, требующих высокой точности при минимальных вычислительных ресурсах — таких как встроенные устройства, интерфейсы взаимодействия человека с компьютером и системы безопасности. Использование наборов данных ORL и HandReader подтверждает применимость предложенного подхода для маломасштабных и реальных сценариев, демонстрируя его потенциал интеграции в ресурсоограниченные системы распознавания. Легковесная структура метода обеспечивает простоту внедрения и доступность, что представляет практическую ценность для организаций и исследователей, работающих в условиях ограничений данных или аппаратных ресурсов. Таким образом, проведенное исследование способствует реализации общей цели — разработке эффективных и устойчивых технологий распознавания, пригодных для разнообразных и ограниченных по данным сред.

Научные гипотезы, основные результаты и их значимость: Центральная гипотеза диссертации заключается в том, что высокая точность распознавания объектов при ограниченных данных может быть достигнута посредством интеграции снижения размерности, адаптивного метрического обучения и регуляризации признаков на основе проекции в единую легкую архитектуру.

Эта гипотеза была подтверждена как аналитически, так и экспериментально с использованием наборов данных ORL и HandReader, что показало: стабильное и точное распознавание возможно не только при больших данных и глубинных архитектурах, но и при оптимальном взаимодействии этапов извлечения признаков и обучения сходства.

Ключевые положения, выносимые на защиту:

1. Разработка PCA–TP (Principal Component Analysis – Triplet Projection) алгоритма и метода, объединяющего линейное снижение размерности с модифицированным метрическим обучением для эффективного и интерпретируемого распознавания объектов.
2. Введение Modified Triplet Similarity Algorithm, интегрирующего косинусное сходство, адаптивный угловой зазор и регуляризацию проекционной согласованности, что усиливает дискриминативную структуру эмбедингов, балансирует внутриклассовую компактность и межклассовое разделение, а также обеспечивает устойчивую сходимость обучения.
3. Экспериментальная проверка предложенного метода и алгоритма, показавшая точность распознавания 90–100% на наборах данных ORL и HandReader при низкой вычислительной сложности и минимальном числе параметров.

Научная новизна исследования заключается в создании единого и легковесного алгоритма и метода распознавания, объединяющего PCA и модифицированный triplet similarity algorithm, включающий cosine similarity, adaptive margin adjustment и projection consistency в единую оптимизационную структуру. Такой подход формирует новое методологическое направление в метрическом обучении для задач с ограниченными данными, сочетая геометрическую регуляризацию и линейное проекционное отображение.

Значимость результатов проявляется как в теоретическом, так и в практическом аспектах. С теоретической точки зрения, исследование углубляет понимание того, как угловое сходство и проекционная регуляризация повышают устойчивость и дискриминативность обучения в условиях малого объема данных. С практической стороны, предложенный алгоритм и метод могут эффективно применяться в ресурсоограниченных или встроенных системах, таких как биометрическая аутентификация, распознавание жестов и интеллектуальный мониторинг.

Публикации и вклад автора: Результаты данной диссертации опубликованы в серии научных статей, которые в совокупности отражают последовательное развитие моего исследования в области распознавания объектов, лиц и жестов в условиях ограниченных данных. Каждая из представленных работ

внесла вклад в формирование предложенного алгоритма PCA–TP и общей методологической структуры диссертации.

“Sign Language Dactyl Recognition Based on Machine Learning Algorithms” (Eastern-European Journal of Enterprise Technologies, 2021, Vol. 4, Issue 2, pp. 58–72) представила систему распознавания казахского дактильного языка с использованием методов машинного обучения. В исследовании были применены алгоритмы Random Forest, Support Vector Machine и Extreme Gradient Boosting к набору данных из 5000 изображений на каждый жест. Достигнута точность классификации 98.86%, 98.68% и 98.54% соответственно. Работа продемонстрировала потенциал классических алгоритмов машинного обучения для высокоточного распознавания жестов. Вклад автора заключался в оказании методической и технической помощи при разработке алгоритмов, программировании и оценке классификаторов, а также в консультировании авторов по вопросам адаптации методологического подхода данного исследования к экспериментальной части диссертации.

“2D Face Recognition Using PCA and Triplet Similarity Embedding” (Bulletin of Electrical Engineering and Informatics, 2023, Vol. 12, Issue 1, pp. 580–586) представила алгоритм PCA–TP, объединяющий уменьшение размерности и триплетное метрическое обучение с проекционным классификатором. Концепция, предварительная обработка и эксперименты были полностью разработаны и реализованы автором. Алгоритм достиг точности до 98.75% на наборе данных ORL при соотношении обучения и тестирования 80/20 и продемонстрировал устойчивость при меньших выборках. Это исследование стало основной экспериментальной базой диссертации, послужив основой для окончательной модели PCA–TP, объединяющей PCA, Modified Triplet Similarity Loss и Projection-based classification.

“PCA and Projection-Based Hand Gesture Recognition” (Journal of Theoretical and Applied Information Technology, 2021, Vol. 99, Issue 15, pp. 3801–3811) была посвящена распознаванию жестов на основе изображений с использованием набора данных HandReader. Автором был предложен, реализован и экспериментально подтвержден метод, сочетающий PCA для сокращения признаков, триплетное встраивание для дискриминативного обучения и проекционную классификацию. При использовании всего 20% данных для обучения и 80% для тестирования алгоритм достиг точности до 95%, значительно превысив результаты методов KNN и MSRC. Исследовательская концепция и анализ полностью выполнены мной, а предложенная методология была впоследствии интегрирована в рамки алгоритма PCA–TP.

“An Enhanced Face Recognition Method for Lighting Problem” (15th International Conference on Electronics, Computer and Computation, ICECCO 2019, pp. 1–4) рассматривала методы предварительной обработки изображений для повышения точности распознавания лиц при изменении

условий освещения. Комбинация Histogram Equalization и Adaptive Single Scale Retinex (ASSR) позволила повысить точность распознавания с 10% (без обработки) до 90%. Вклад автора заключался во внедрении PCA и проведении экспериментальных тестов. Результаты этой работы использовались при разработке методологии предварительной обработки данных для задач распознавания лиц в диссертации.

“Gesture Recognition Based on Sparse Reconstruction” (14th International Conference on Electronics, Computer and Computation, ICECCO 2018, pp. 206–212) исследовала применение алгоритма MSRC для распознавания жестов и его сравнение с классификаторами KNN и Random Forest. Все эксперименты и реализация алгоритмов были полностью выполнены автором. Результаты показали, что MSRC обеспечивает более высокую точность распознавания (до 97.5%) и меньшую вычислительную сложность на наборе данных HandReader. Данное исследование напрямую повлияло на выбор направления диссертации — объединение методов разреженного представления и проекционного обучения при ограниченных данных.

“Comparative Analysis of Recognition Algorithms for Hand Gestures on the Basis of Various Representations of Images” (Herald of the Kazakh-British Technical University, 2019, Vol. 16, Issue 1, pp. 50–54) представила сравнительный анализ алгоритмов KNN, Logistic Regression и Decision Tree при использовании признаков на основе raw pixel и гистограмм. Все эксперименты и анализ были проведены автором. Классификатор Logit с признаками raw pixel показал точность 98%, но требовал больше времени вычислений. Работа подтвердила влияние способа представления данных на производительность алгоритмов и послужила стимулом для последующего использования PCA для компактного кодирования признаков.

“Information Technology for Sign Language Alphabet Units Modeling and Recognition” (Bulletin of KazNITU, 2020, Vol. 5, pp. 364–371) представила кроссплатформенную технологию моделирования и распознавания жестов, основанную на виртуальных моделях руки и стандартизированных протоколах YAML и PostgreSQL для баз данных жестов. Система поддерживала мультязычные наборы жестов и продемонстрировала эффективность кроссплатформенных архитектур. Вклад автора включал разработку компонента распознавания на основе CNN и участие в системном проектировании.

“Comparative Analysis of Face Recognition Algorithms in the Problem of Visual Identification” (Suleyman Demirel University Bulletin: Natural and Technical Sciences, 2021, Vol. 54, Issue 1, pp. 61–67) проанализировала классические алгоритмы, такие как EigenFace и Gabor фильтры, для задач визуальной идентификации. Автор руководил проектом, помогая магистрантам формировать методологию исследования, подготавливать данные и реализовывать эксперименты. Хотя результаты данной работы не

использовались напрямую в диссертации, они позволили сформировать сравнительный контекст между классическими методами и алгоритмом PCA-TP.

Наконец, “**Methods for Handling Frames of Hand Gesture Recognition**” (Suleyman Demirel University Bulletin: Natural and Technical Sciences, 2021, Vol. 54, Issue 1, pp. 45–51) рассматривала методы обработки статических изображений жестов, извлеченных из видеопотоков, включая сглаживание, бинаризацию, поворот и изменение фона на основе модели цвета кожи. Автор осуществлял научное руководство и контроль за экспериментальными процедурами, обеспечивая методологическую точность. Представленные в работе методы предварительной обработки были использованы при подготовке данных для экспериментов в диссертации.

Эти девять публикаций отражают последовательное развитие научного направления, завершившееся созданием предложенного алгоритма PCA-TP и его методологической базы. Вклад автора охватывает полный цикл научного исследования — от разработки идеи и алгоритма до проведения экспериментов и публикации результатов, что привело к созданию эффективной, интерпретируемой и компактной системы распознавания в условиях ограниченных данных.

Объем и структура диссертации: Диссертация состоит из 94 страниц и включает несколько разделов, которые всесторонне охватывают как теоретические основы, так и экспериментальную проверку проведённого исследования. Работа начинается с введения, за которым следует обширный обзор литературы, где рассматриваются история, актуальность и основные проблемы распознавания образов, с особым вниманием к условиям ограниченных наборов данных и современным подходам, таким как data augmentation, GANs, transfer learning и few-shot learning. В разделе методологии исследования подробно описаны используемые наборы данных (включая HandReader и ORL), применённые алгоритмы (PCA, модифицированная функция triplet loss, projection) и структура предложенного алгоритма PCA-TP. В разделе эксперименты и результаты представлена серия экспериментов по распознаванию жестов и лиц, включая предварительную обработку данных, оценку классификаторов и сравнительный анализ различных методов. Диссертация содержит 156 библиографических источников, что подтверждает прочную научную и академическую основу работы. Для повышения наглядности и более глубокого понимания сложных процессов в тексте приведены 31 диаграмма и 18 таблиц. Такая структурная организация обеспечивает логичное и последовательное изложение материала, сосредоточенного на разработке эффективного алгоритма распознавания объектов в условиях ограниченных

данных, основным научным вкладом которого является предложенный метод PCA–TP.