Innovative Teaching and Learning Vol. 5, No. 1, pp. 19-34
http://doi.org/10.4208/itl.20230102 June 2023

Teaching Exploration of Pattern Analytics Course-Bridging the Distance Between Theory and Practice

Hongjian Shi*, Chengyan Lin, and Songze Zhang

BNU-HKBU United International College

Abstract

This article is an exploration of the teaching and optimization reform project of computer technology application courses. According to the teaching status of a strong theoretical course, this paper takes the course of Pattern Analytics as an example to explore how to enhance students' understanding of theoretical knowledge, enthusiasm, and realization ability by increasing practice of theory-related practical projects in the real world.

Keywords: strong theoretical course, pattern analytics, increasing practice

*Corresponding author: Hongjian Shi. E-mail: shihj@uic.edu.cn

Introduction

Pattern Analytics is the first critical processing step taken by all intelligent systems to extract meaningful information from complex sensor data. Pattern Analytics comes from statistics and focuses on the study of data classification methods (Theodoridis & Koutroumbas, 2006). Pattern Analytics is inextricably linked to many other computer research fields, such as machine learning (Bishop, 2006), image processing (Gonzalez & Woods, 2009), computer vision (Zhang, 2016), etc., and is closely related to many high-tech fields such as artificial intelligence, video and image identity Identification and medical image-aided diagnosis, etc. (Richard. et al., 2000). Therefore, Pattern Analytics has become a compulsory course for graduate students majoring in computer science in most colleges and universities internationally. The teaching quality of the courses closely related to high-tech industries, represented by Pattern Analytics, directly affects the quality of training high-tech talents and determines whether these high-tech talents could quickly adapt to industrial production and apply theories to practice in industries. The main content of this paper is to reform the traditional teaching mode of the Pattern Analytics course by transforming this almost purely theoretical course into one theory and practice balanced course and increasing practical projects close to high-tech industrial production. This largely increases students' enthusiasm for solving practical problems and then enhances their understanding abilities to course theories.

Analysis of the Present Situation of Teaching Pattern Analytics Course

There are many problems in traditional course teaching of Pattern Analytics (Liu. et al., 2011). First of all, the traditional teaching of the Pattern Analytics course is mostly theoretical. The methods of Pattern Analytics involve a large number of statistical principles and mathematical derivations. The teaching of the Pattern Analytics course can easily become the teaching of one mathematics course. It is filled with a large number of mathematical formulas and the content is abstract, which makes it difficult for students to understand and may make them feel daunted. Second, the proportion of experiments in traditional Pattern Analytics courses is very low. Students generally regard the experimental class as extra time for the theoretical class, and couldn't complete the "implementation" of the theory. Over time, this course has become a de facto theoretical course, and the implementation is only a little bit. Third, the traditional Pattern Analytics course lacks the content of the multidisciplinary applications and is more likely a purely theoretical statistics course. Even if students have a thorough understanding of the theory, they still feel at a loss when they encounter practical applications without the inspiration of drawing inferences from one instance.

The Scholarship of Teaching and Learning (SoTL) points that effective assessment of student learning is necessary. Assessment is not only a way to measure student achievement, but also a way to enhance student learning by providing feedback, motivation, and guidance (Open University, 2021). Assessment can also inform teaching decisions and improve instructional design. However, assessment in computer science education is not trivial, as it involves measuring complex cognitive processes, such as problem-solving, debugging, code comprehension, and program design. Moreover, assessment in computer science education needs to balance between validity, reliability, fairness, efficiency, and scalability (Sweller, 2016). In computer science, programming assignments are created with the goal of showcasing learning objectives and engaging students (Layman. et al., 2007). These assignments can involve mathematical functions, business applications, or game developments. Researchers have explored ways to motivate students through choices in classroom assignments (Aycock & Uhl, 2005; Cliburn & Miller, 2008; Layman et al., 2007). Such choices include completing additional problems, adding embellishments to existing assignments, or optionally resubmitting corrected assignments (Becker, 2006).

To change students' stereotypes of the Pattern Analytics course and improve students' learning enthusiasm and theoretical application ability, we continue to explore the teaching of the Pattern Analytics course, increase the number of theoretical application projects, apply the theory into practice, and make mathematical formulas more vivid and specific in applications.

Similar to the SoTL that discovers new teaching problems and teaching research topics in teaching interaction, our teaching exploration guides students to discover their research interests and topics related to the curriculum through the interaction of theory and practice.

Teaching Exploration of Pattern Analytics Course Curriculum Provision

Course Intended Learning Outcomes (CILOs)

The following are the intended learning outcomes written clearly in our syllabus in Table 1. Our assessments of intended learning outcomes are through assignments and laboratory classes that occupy half of the assessments of the entire course, connecting theory with practice.

Table 1

Session CILOs

CILO	By the end of the course, students should be able to
CILO 1	Derive and implement Bayesian Theory
CILO 2	Perform Parametric and Non-Parametric Density Estimation
CILO 3	Design Linear Classifiers for separable and non-separable pattern
CILO 4	Implement classifiers using neural networks

Bayesian Theory is used to predict a certain probability based on some limited past data. For example, using limited information (past weather data) to predict the probability of rain tomorrow. The underlying idea of the Bayesian Theory is: the newly observed sample information will revise people's prior knowledge of things. It's like human beings have only a little prior knowledge of nature at the beginning. As more samples are obtained through continuous observation and practice, people's understanding of the laws of nature becomes more and more thorough.

In the Bayesian model, each parameter can be regarded as a random variable, that is, it follows a certain distribution. Then, we will be able to introduce our prior knowledge of the data by assuming the form of a parametric distribution. Parametric estimation assumes that the research problem can be modeled using a normal distribution or a binomial distribution, then we can use known training data to estimate the parameters in the distribution assumed.

The parameter estimation is performed on the parameters using the overall information assumption, but this assumption is sometimes not true. Non-parametric estimation directly uses the prior knowledge of known categories of training samples to directly perform statistical testing and judgement analysis without considering the original overall distribution, or without making assumptions about parameters.

Pattern analytics (or recognition) is to find an effective classification decision rule based on the cognition of the characteristics of a class of objects, which can correctly classify new samples. For example, the known sample set is divided into two categories. If a boundary between categories can be found in the feature space, it can be determined which category the sample belongs to by judging which side of the boundary the sample being identified is located on. This boundary can be called the "classification decision boundary." The discriminant function corresponding to the classification decision boundary can be linear or non-linear, also separable or non-separable.

A neural network is a mathematical model or computational model that imitates the structure and function of a biological neural network. Neural networks are computed by a large number of artificial neural connections. In most cases, the artificial neural network can change the internal structure on the basis of external information, and it is an adaptive system, that is, it has the function of learning. Neural networks can be applied to many aspects of pattern classification, such as parameter estimation.

Course Hours Arrangement

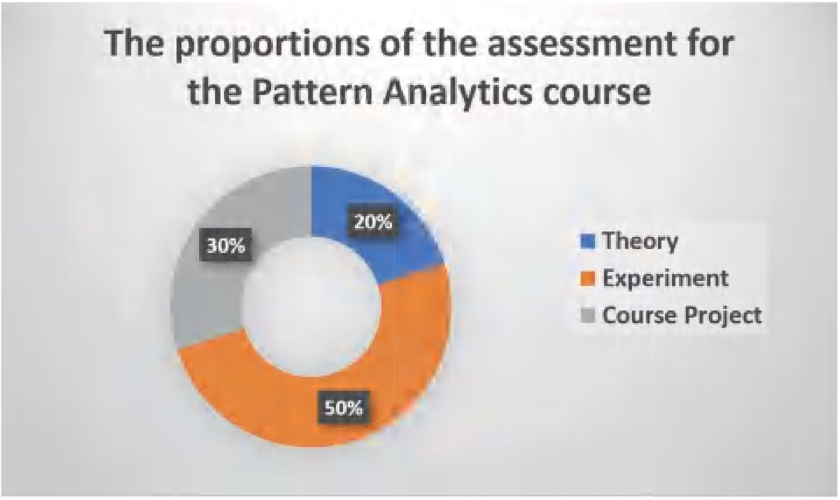
The first step in the reform of the Pattern Analytics course is to increase the weight of the experimental part in the arrangement of class hours. The traditional arrangement of the Pattern Analytics course generally includes 3 semester hours per week of teaching theories. Some colleges and universities will arrange several semester hours of one separate experimental course. At our university, I arranged 2 hours of teaching theories followed by 1 hour of teaching code implementation of the lectured theories and then handy coding by students in class. In particular, our university, i.e., Beijing Normal University-Hong Kong Baptist University United International College is a jointly funded university of higher education by Beijing Normal University and Hong Kong Baptist University. Lecturers can arrange 1 additional hour of tutorial class per week according to teaching needs, and the teaching assistants will answer questions for students. To improve the teaching effect of the experimental class, the tutorial class immediately follows after the experimental class. Students can continue to do their unfinished implementation of projects in the tutorial class or after class. In this way, students can discuss the problems encountered in both lecture and experimental classes with the teacher and teaching assistants. As the instructor of the Pattern Analytics course, I have nearly 20 years of working experience in industries and more than 10 years of teaching experience. The teaching assistants and I have been trying to answer students' theoretical and implementational questions and help them recognize the concepts, theories, and code implementation.

Assessment Settings

The assessment of this course is divided into three parts: theory via assignments, experiment via code implementation, and course project via teamwork to solve complicated real-world problems. The proportion of each part of the assessment is shown in Figure 1.

Figure 1

The Proportions of the Assessment for the Pattern Analytics Course



Content-related assignments are for theory understanding, and experiments are designed for code implementation of algorithms and formulas. Assignments are mathematical derivation or logical reasoning in theoretical environments closely following lectures in class. Students are required not only to master the derivation of important theoretical knowledge but also to learn some current theories that are not included or detailed in the textbook.

The contents of the experimental portion can be roughly divided into two parts, one being the code implementation of the theoretical knowledge taught in the theoretical lectures, and the other being the algorithm applications of Pattern Analytics. Due to the different subject backgrounds at the undergraduate stage, the course does not limit the programming language for code implementation. Students can use their familiar programming languages to complete the experiments of code implementation. Considering that graduate students will need to engage in academic or professional computer science research in the future, during the teaching process, the lecturer uses MATLAB to demonstrate the applications of the theory and also teaches students how to use MATLAB to write codes.

The course project is a comprehensive assessment of the students' abilities to apply what they have learnt in class to meaningful real-world problems. When half of the course content is taught, the lecturer will publish a list of course project topics encountered in our daily life or industries. Each student or two students as a team can choose a topic from the list or a topic related to the course defined by the team. The self-defined projects must be approved by the lecturer to make sure that the projects are soundly defined and meaningful in the real world or industries. Students need to finish the course projects as completing scientific research projects. They need to survey the research or potential product progress on the projects, existing methods, and their results. They need to conduct implementation, compare the existing methods with their corresponding results, and attempt to improve or invent new solutions to the selected projects. The course project has higher requirements for students, and students need to use the theories learned in class to solve practical research or real-world problems.

Curriculum Content Design-Strengthening the Coherence and Logic of the Course Contents

Teaching a course well is like telling a good story: A clever opening draws the reader in; the plot is interlocking and fascinating; the end leaves suspense and inspires thinking. The contents of the Pattern Analytics course start with a practical problem of bass and salmon classification, leading to the basic concepts and ideas of Bayesian decision-making, extending from binary classification problems to multi-classification problems, estimating from known model parameters to unknown parameters. The course contents are from shallow to deep, with good coherence and logic, which can help students better understand the contents of the course and draw the knowledge context and knowledge process of the course.

Table 2 introduces the content arrangement of the course to show the coherence and logic of the course contents.

Table 2

Pattern Analytics Course Contents

Title	Contents
Course Introduction	Introduce the course schedule, assessment rubrics, and course objectives. What is the pattern? What is pattern analytics? Introduce the classification of bass and salmon in production and life, understand the general process of Pattern Analytics, discuss the solution to this problem, and pave the way for the introduction of Bayesian decision.
Bayesian Decision Rule (Bayes & Price, 1763)	Taking the classification of bass and salmon as an example, introduce the concepts of prior probability, posterior probability, likelihood, and probability error of Bayesian decision. Introduce the Bayesian Decision. Introduce the minimum risk and discriminant functions. Introduce Gaussian distribution. On the premise that the samples obey the Gaussian distribution, analyze the expression of the discriminant function under different distribution parameters.
Mathematical Foundation	Introduce the knowledge of probability and statistics required for the course, including probability, random variable, probability density function, covariance, normal distribution, etc. Introduce the linear algebra knowledge required for the course, including basic matrix operations, inner product, determinant, trace, inverse matrix, matrix derivation, Taylor expansion, eigenvalue, eigenvector, etc. This chapter supplements and reviews the mathematical foundations that students need to master in the course study, and lays a theoretical foundation for the study of subsequent courses.

Title	Contents
Maximum-Likelihood and Bayesian Parameter Estimation (Braverman, 1962)	The prior probability and likelihood in Bayesian decision are usually unknown in practical applications. Therefore, we need to use the existing samples to estimate the required parameters. The process of parameter estimation is the process of training a classification model. This chapter focuses on the idea, theoretical derivation, and algorithm steps of maximum likelihood estimation and Bayesian estimation.
Bayesian Estimation–PCA and MDA (Abdi & Williams, 2010)	Classification in the real world is usually multi-dimensional (multi-features), and the dimensionality of training samples and the number of samples greatly increase the computational complexity of parameter training, which naturally increases the training time. For the model to complete parameter training efficiently and accurately, it is necessary to reduce the dimensionality of the training samples. The principal component analysis can express data with a small number of features, and multiple discriminant analysis can best separate data in the least squares sense. The combination of principal component analysis and multiple discriminant analysis can realize a fast and efficient training of the Bayesian classification model.
Expectation-Maximization and Mixture Density Estimation (McLachlan & Krishnan, 2007)	This chapter aims at solving real-world problems. Real-world training samples may be incomplete, and some samples are missing for certain features. In the case of incomplete sample data, how to estimate the distribution of samples, maximum expectation and mixed density estimation provide solutions. Up to this chapter, students have been able to use what they have learned to train the parameters of Bayesian decision on samples to solve most classification problems in the real world.

Title	Contents
Parametric Models–Hidden Markov Models (Eddy, 2004)	This chapter discusses methods for classifying samples related to time series. Samples for timing-related classification problems, such as sounds, DNA sequences, etc., whose features transition between different states over time. Hidden Markov models can solve such classification problems well.
Parzen Window Method (Parzen, 1962)	Parzen Window Method (Parzen, 1962) The parameter estimation methods introduced above are all parametric methods suitable for samples with unimodal distribution. The window method introduced in this chapter is a non-parametric method, which can estimate the likelihood and prior probability of samples with any distribution.
Probabilistic Neural Networks (Mohebali et al., 2020)	This chapter is an extension of the previous chapter. Parzen Window is also called a probabilistic neural network. It is a neural network model based on statistical principles. Its theoretical derivation is related to the Bayesian decision theory. It can be regarded as a classification model based on the Bayesian decision. By this chapter, students can feel the development of the classification model from statistics to Pattern Analytics, and then to the development and close connection between neural networks.
Linear Discriminant Analysis (Izenman, 2013, pp. 237–280)	Most of the models learned in the previous chapters use training samples to estimate the parameters of the probability density when the probability density is known or given. The method introduced in this chapter is similar to the window method, for the cases where the correct forms of the discriminant functions are known but the probability density functions are unknown.

Title	Contents
An Algorithm-Independent Machine Learning	For this chapter, lectures will give a selective introduction of an algorithm-independent machine learning to the students, such as famous no free lunch algorithm and ugly duckling algorithm.

If the context of knowledge is not clear, students would not know why. This is a problem faced by many theoretical courses in teaching. Strengthening the coherence and logic of the course contents can help students better grasp the knowledge context of the course and look at the problems more macroscopically. Especially for graduate students, the coherence of knowledge and the openness of thinking can help them go further on the road to their future scientific research or industrial projects in real world.

Coursework Design

Pattern Analytics is developed from statistics and is widely used in the real world. This course requires students not only to have a solid grasp of the relevant theoretical background and derivation of the course contents, but also to apply theoretical knowledge and algorithms to solve real classification problems. Therefore, the assignments for this course are divided into theoretical and experimental parts. The theoretical part helps students to understand and reinforce the theoretical knowledge taught in the course. For some problems that have not been fully developed in class or are not detailed in the textbooks, students are asked to supplement and learn in the form of topics in theoretical assignments. The experimental part requires students to use programming languages to implement the theoretical knowledge and algorithms they have learned, which is not only the consolidation of theoretical knowledge but also the concretization of abstract knowledge. In the process of practice, students can verify the correctness of their understanding of theory. In the process of continuously transforming theory into practice, they can accumulate certain results with transformation experience. By insisting on completing this course, students can achieve a qualitative leap in theoretical understanding and code practice in high-intensity, fast-paced theoretical learning, and theoretical practice. Table 3 shows the distribution of theoretical and experimental assignments for each chapter of the course contents. The experimental assignments or labs are designed to implement theoretical concepts and algorithms lectured in classes. Most concepts and algorithms require students do their implementations from simple random number generation, Gaussian distribution, and central limit theorem, to the Bayesian decision, Bayesian estimation, Parzen window method, hidden Markov chain, Expectation-Maximum algorithm, and neural networks.

Table 3

Schema of Theoretical Assignments and Practical Labs

Title	Theoretical Assignment	Lab Assignment
Introduction		
Bayesian Decision Rule	√	√
Mathematical Foundation	√	√
Maximum-Likelihood and Bayesian Parameter Estimation	√	√
Bayesian Estimation – PCA and MDA	√	√
Expectation -Maximization and Mixture Density Estimation	√	√
Parametric Models – Hidden Markov Models	√	√
Parzen Window Method	√	√
Probabilistic Neural Networks	√	√
Linear Discriminant Analysis	√	
An Algorithm-Independent Machine Learning		

In consideration of the efforts on implementation, the final examination is concluded with a final course project to solve real-world problems. Students can choose a topic from the list of course project topics provided by the teacher, or they can identify a topic of interest for research. Table 4 lists some of the students' course project topics based on some potential new research problems derived from this course or the combination with other skills from other courses.

Table 4

List of Course Project Topics

No.	Title
1	Tooth segmentation in CBCT images via deep learning
2	Mask2Vote: Mask to Voting-based Ensemble Segmentation for Medical Image
3	An End-to-End Edge-Cloud Co-Inference Framework based on PCA
4	Image Segmentation based on Bayesian Rule
5	Object Tracking based on MDP
6	Least Square for Edge Closing and Image Enhancement
7	Face Recognition System with PCA
8	Classify Random Data by K-Means and K-NN
9	Naïve Bayesian Classification for Sentiment Analysis
10	Use HMM for Fingerprint Image Segmentation
11	Use PCA to Speed Up Optical Character Recognition
12	License Plate Location Recognition based on Threshold Segmentation

Student Feedback for Pattern Analytics

Pattern Analytics is an elective course for graduate students. In addition to Pattern Analytics, another elective course called Digital Image Processing is also involved in the teaching and optimization reform project. Students who worked well with Pattern Analytics will then take Digital Image Processing to improve their knowledge-application transfer capability. Last year, about 13.64% of students in Pattern Analytics took Digital Image Processing. This year, that percentage reached about 28.57%. This means our teaching mode is accepted gradually by more and more students.

This course in the name of Pattern Recognition was also taught by the first author at Southern University of Science and Technology and received high assessments from students and peer colleagues. For example, in the Spring of 2017, the first author taught the Pattern Recognition course and won 98 out of 100 from peer colleagues' judgement of teaching and 97.65 out of 100 from students. Many students felt that they learned and practiced a lot from this course. Students talked with the first author and said that this course is a real course to help them master the skills. Therefore, the course became one of the three essential courses for graduate students at the Department of Electric and Electronics, and also the selective course for the Department of Computer Science, the Department of Biomedical Engineering at Southern University of Science and Technology in 2018. Some undergraduate students went to their PhD study in the field of Pattern Recognition after graduation.

Conclusion

This paper explains how to break the traditional rigid teaching mode of the Pattern Analytics course from the four aspects: course hours, assessment, curriculum contents, and course work. Students can understand better the course contents in theoretical background and theoretical derivation by improving the coherence and logic of the course contents. Students' emphasis on the theoretical applications will be enhanced by optimizing course contents, experimental practice, class hours, and assessment settings. The course design of both theories and implementations improves students' theoretical levels and practical implementation abilities to solve real-world problems. Theory and practice conjoin for students.

Acknowledgments

This work was supported by research grants of BNU-HKBU United International College (R202114, UICR0400054-21CL, UICR0500005-21), Guangdong Higher Education Key Platform and Research Project (2020ZDZX3039), Guangdong Provincial Key Laboratory of Interdisciplinary Research and Application for Data Science (2022B1212010006), and Guangdong Higher Education Upgrading Plan (2021-2025) (R0400001-22).

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics, 2*(4), 433–459. https://doi.org/10.1002/wics.101

Aycock, J., & Uhl, J. (2005). Choice in the classroom. *Association for Computing Machinery Special Interest Group on Computer Science Education Bulletin, 37*(4), 84–88. https://doi.org/10.1145/1113847.1113883

Bayes, T., & Price, R. (1763). An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions (1683-1775), 53*(1763), 370–418. https://www.jstor.org/stable/105741

Becker, K. (2006). How much choice is too much? *SIGCSE Bulletin, 38*(4), 78–82. https://doi.org/10.1145/1189215.1189176

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Braverman, D. (1962). Learning filters for optimum pattern recognition. *IRE Transactions on Information Theory, 8*(4), 280–285. https://doi.org/10.1109/TIT.1962.1057722

Cliburn, D. C., & Miller, S. (2008). Games, stories, or something more traditional: The types of assignments college students prefer. In S. Rodger, J. Dougherty, M. Guzdial, S. Fitzgerald, & E. Walker (Eds.), *Proceedings of the 39th Special Interest Group on Computer Science Education Technical Symposium on Computer Science Education* (pp. 138–142). Association for Computing Machinery. https://doi.org/10.1145/1352135.1352184

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification (2*nd *ed.)*. Wiley-Interscience.

Eddy, S. R. (2004). What is a hidden Markov model? *Nature Biotechnology, 22*(10), 1315–1316. https://doi.org/10.1038/nbt1004-1315

Gonzalez, R. C., & Woods, R. E. (2009). *Digital image processing* (3rd ed.). Pearson Education India.

Izenman, A. J. (2013). *Modern multivariate statistical techniques: Regression, classification, and manifold learning.* Springer. https://doi.org/10.1007/978-0-387-78189-1_8

Layman, L., Williams, L., & Slaten, K. (2007). Note to self: make assignments meaningful. *ACM SIGCSE Bulletin, 39*(1), 459–463. https://doi.org/10.1145/1227504.1227466.

Liu, D., Yuan, Z., & Zhao, J. (2011). *Information, Communication and Education Application (ICEA 2011)*(pp.253-258). Information Engineering Research Institute.

McLachlan, G. J., & Krishnan, T. (2007). *The EM algorithm and extensions.* John Wiley &

Sons. https://doi.org/10.1002/9780470191613

Mohebali, B., Tahmassebi, A., Meyer-Baese, A., & Gandomi, A. H. (2020). Probabilistic neural networks: a brief overview of theory, implementation, and application. In P. Samui, D. T. Bui, S. Chakraborty, & R. C. Deo (Eds.), *Handbook of Probabilistic Models* (pp. 347–367). Butterworth-Heinemann. https://doi.org/10.1016/B978-0-12-816514-0.00014-X

Open University. (2021, August). *Scholarship of teaching and learning in STEM*. OpenLearn. http://bit.ly/47Heuem

Parzen, E. (1962, September). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics, 33*(3), 1065–1076. https://doi.org/10.1214/AOMS/1177704472

Sweller, J. (2016). Cognitive load theory and computer science education. *In Proceedings of the 47th ACM Technical Symposium on Computing Science Education (SIGCSE '16).* New York, NY, USA: Association for Computing Machinery. https://doi. org/10.1145/2839509.2844549

Theodoridis, S., & Koutroumbas, K. (2006). *Pattern recognition* (3rd ed.). Elsevier. https://doi.org/10.1016/B978-0-12-369531-4.X5000-8

Zheng, N., Loizou, G., Jiang, X., Lan, X., & Li, X. (2007). Computer vision and pattern recognition. *International Journal of Computer Mathematics, 84*(9), 1265–1266. doi:10.1080/00207160701303912

Received January 28, 2023
Accepted April 24, 2023