

GIÁM SÁT AN NINH DÙNG KỸ THUẬT XỬ LÝ HÌNH ẢNH

SECURITY SURVEILLANCE WITH IMAGE PROCESSING TECHNIQUES

Hồ Quốc Thiên¹, Lê Mỹ Hà²
Trường đại học Sư phạm Kỹ thuật TP.HCM

TÓM TẮT

Phát hiện người và nhận dạng hành động đóng vai trò quan trọng trong nhiều ứng dụng giám sát an ninh cũng như trong các bộ mô tả cần khai thác mối tương quan giữa đối tượng và ngữ cảnh trong video. Bài báo hướng đến việc phát hiện và nhận dạng hành động người trong một khu vực giám sát, thực thi trên chương trình MATLAB. Kết quả đạt được của đề tài: phát hiện thành công người và nhận dạng các loại hành động đã huấn luyện như đi, chạy, lò cò, trượt ngang, nhảy. Đề tài sử dụng thuật toán kênh tổng hợp đặc trưng (ACF) huấn luyện trên tập mẫu Caltech để xây dựng bộ phát hiện người – Tracking dùng bộ lọc kalman – và dùng 5 đặc trưng HOG sắp xếp liên tiếp nhau trên miền thời gian để phát hiện được hành động của đối tượng thông qua bộ phân loại SVM.

Từ khóa: ACF, tracking, kalman, HOG, SVM

ABSTRACT

Human detection and action recognition are important role in many applications such as security surveillance or also describe the relationship between objects and scenes in the video. The article aims to detect human and identify the action of them in a regional monitoring in MATLAB environment. The experimental results have detected human and recognized 5 actions (trained): walk, run, skip, side, jump. This project used ACF (Aggregated Channel Features) training in Caltech dataset to create the human detector – Tracking with Kalman filter -and arrange 5-HOG consecutive features to make an actions descriptor through SVM classifier.

Keywords: ACF, tracking, kalman, HOG, SVM

1. GIỚI THIỆU

Trong cuộc sống hàng ngày, con người luôn không ngừng vận động và phát triển. Mọi thứ xung quanh chúng ta ngày một hiện đại hơn, phát triển hơn. Do đó việc có những thiết bị thông minh hiểu được hành động người trong video có thể tạo thuận lợi hơn trong việc tương tác giữa người và máy, và ngoài ra có thể dựa vào những hành động mà máy tính hiểu ứng dụng cho các camera giám sát trong các khu vực bất kì. Để các thiết bị có thể nhận dạng được hành động của người gồm 3 phần: Phát hiện người – tracking theo mục tiêu – nhận dạng hành động.

Phát hiện đối tượng (người): trong thị giác máy tính là công việc tìm kiếm các vật thể trong ảnh hoặc trong video.

Bám mục tiêu (tracking): Sau khi phát hiện được người trong ảnh, việc bám mục tiêu (tracking) để theo dõi các đối tượng cũng là yếu tố cần thiết để có thể xác định được hành động của người đó trong nhiều khung ảnh

liên tiếp hoặc video.

Nhận dạng hành động người: dùng bộ phân loại để nhận dạng được hành động của đối tượng.

2. TÌNH HÌNH NGHIÊN CỨU

Phát hiện người là hướng nghiên cứu dành được rất nhiều sự quan tâm vì khả năng ứng dụng của nó. Manoranjan Paul, [1] đã trình bày tổng quan (review) về hầu hết các phương pháp phát hiện người cũng như các ứng dụng của nó trong thực tế. Ngoài ra, N.Dalal; B.Triggs [2] đề xuất đặc trưng HOG và phân loại SVM dùng cửa sổ trượt để phân loại người và nền (background) trong khung ảnh, là một phương pháp nền tảng và được được biết đến rộng rãi trong lĩnh vực phát hiện người. Viola-Jone[3] đề xuất mô hình phân tầng để nhận dạng đối tượng.

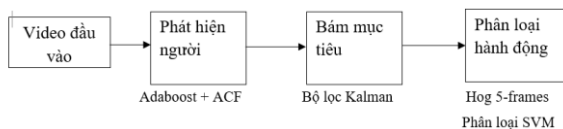
Bám mục tiêu (tracking) giúp liên kết các đối tượng mục tiêu trong khung hình video liên tục. Một vài kỹ thuật bám mục tiêu (tracking) đã được nghiên cứu như CAMshift[4], bộ

bám mục tiêu (tracker) KLT[5], dùng superpixel để bám mục tiêu (tracking) [6] kết hợp với KNN để phân loại màu cho kết quả tốt nhưng chi phí tính toán cao.

Với nhận dạng hành động Laptev và các đồng nghiệp đã giới thiệu bộ miêu tả mà khi sắp xếp các đặc trưng HoG (Histogram of Oriented Gradients) – HoF (Histogram of Optical Flow) trên miền thời gian có thể rút trích thông tin về dáng vóc và chuyển động, tác giả còn giới thiệu một đặc trưng STIP (Space-Time Interest Points) để miêu tả hành động trong không gian cục bộ [7]. Nazli Ikizler-Cinbis và Stan Sclaroff[8] đề xuất phương pháp cho nhận diện hành động người mà kết hợp nhiều đặc trưng của các thực thể như: Đối tượng, khung cảnh và con người. Lê Việt Tuấn[9] phân tích tổng hợp về các phương thức nhận diện hành động trong video bằng cách kết hợp các phương pháp.

3. PHƯƠNG PHÁP THỰC HIỆN

Mục tiêu của đề tài là xây dựng chương trình phát hiện và bám theo đối tượng là người trong video thu được từ camera, từ đó phân loại hành động của đối tượng người trong video. Phương pháp thực hiện theo sơ đồ sau:



Hình 1: Lưu đồ của hệ thống

Dữ liệu là video thu được từ camera sẽ được xử lý để phát hiện có người trong khung ảnh hay không bằng cách dùng thuật toán ACF và AdaBoost.

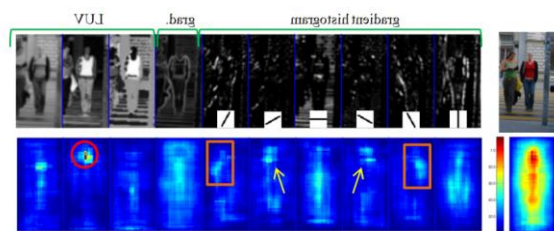
Dùng bộ lọc Kalman để bám mục tiêu là người trong khung ảnh dựa vào vị trí trọng tâm của đối tượng trong khung ảnh

Sử dụng đặc trưng HOG để tính đặc trưng cho đối tượng chuyển động trong các khung ảnh liên tiếp (5 frame), kết hợp với thuật toán phân loại SVM để phân loại hành động của đối tượng.

3.1. Phát hiện người dùng đặc trưng ACF kết hợp với phân loại Adaboost

Đặc trưng kênh tổng hợp (ACF) được đề xuất bởi Dollar[10] sử dụng việc tra cứu

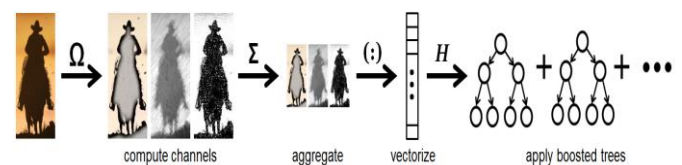
điểm ảnh trong các kênh tổng hợp để giảm thời gian rút trích đặc trưng mà không cần xây dựng các ảnh tích phân. Mỗi đặc trưng ACF đại diện cho một khối mẫu cục bộ, làm đơn giản hơn và tính toán nhanh hơn. Hiệu quả phối hợp của các kênh đặc trưng và việc chiến lược lựa chọn đặc trưng bằng cách xếp chồng các bộ phân loại làm các khung cửa sổ ACF hoạt động tốt hơn hầu hết các phương pháp khác về cả độ chính xác và tốc độ[11]. Các kênh trong ACF được sắp xếp bằng phương pháp ngoại suy để thu được kết quả mong muốn và giảm thời gian tính toán.



Hình 2: Các kênh đặc trưng trong ACF

ACF dùng các kênh: chuẩn hóa biên độ gradient, HOG (6 kênh) và các kênh hệ màu LUV. Trước khi tính toán 10 kênh này, I được làm mịn với một bộ lọc $[1 \ 2 \ 1]/4$. Các kênh được chia thành các khối 4x4 và các điểm ảnh trong mỗi block được tính tổng.

Bộ phát hiện: để phát hiện khách bộ hành, Adaboost được dùng để huấn luyện và phối hợp 2048 quyết định trên 128.64.10/16 = 5120 đặc trưng ứng cử (kênh tra cứu đặc trưng) trong mỗi cửa sổ 128x64.



Hình 3: Tổng quan về bộ phát hiện ACF.

Phương pháp AdaBoost

AdaBoost (Freund & Schapire, 1995) là một bộ phân loại mạnh phi tuyến phức, hoạt động trên nguyên tắc kết hợp tuyến tính các bộ phân loại yếu để tạo nên một bộ phân loại mạnh. AdaBoost sử dụng trọng số để đánh dấu các mẫu khó nhận dạng. Trong quá trình huấn luyện cứ mỗi bộ phân loại yếu được xây dựng thì thuật toán sẽ tiến hành cập

nhập lại trọng số để chuẩn bị cho việc xây dựng bộ phân loại tiếp theo. Cập nhật bằng cách tăng trọng số của các mẫu nhận dạng sai và giảm trọng số của các mẫu được nhận dạng đúng bởi bộ phân loại yếu vừa xây dựng. Bằng cách này thì bộ phân loại sau có thể tập trung vào các mẫu mà bộ phân loại trước nó làm chưa tốt. Cuối cùng các bộ phân loại yếu sẽ được kết hợp lại tùy theo mức độ tốt của chúng để tạo nên một bộ phân loại mạnh.

Bộ phân loại yếu h_k được biểu diễn như sau:

$$h_k(x) = \begin{cases} 1 & \text{nếu } p_k f_k(x) < p_k \theta_k \\ 0 & \text{nếu ngược lại} \end{cases} \quad (3.1)$$

3.2. Tracking – bám mục tiêu dùng bộ lọc Kalman

Bộ lọc Kalman là một tập hợp các phương trình toán học mô tả một phương pháp tính toán truy hồi hiệu quả cho phép ước đoán trạng thái của một quá trình (process) ^{thời gian} sao cho trung bình phương sai của độ lệch (giữa giá trị thực và giá trị ước đoán) là nhỏ nhất.

Bộ lọc Kalman ước lượng trạng thái của một quá trình được mô hình hóa một cách rời rạc theo thời gian bằng một phương trình ngẫu nhiên tuyến tính như sau:

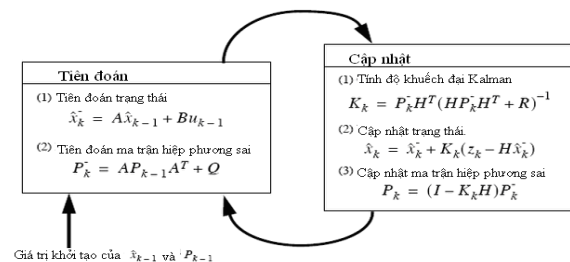
$$x_{k+1} = Ax_k + Bu_k + w_k \quad (3.2)$$

Với giá trị đo lường $z \in \mathcal{R}^n$

$$z_{k+1} = Hx_k + v_k \quad (3.3)$$

Thuật toán Kalman bao gồm 2 bước :

- 1- Ước đoán trạng thái tiên nghiệm, và sau đó,
- 2- Dựa vào kết quả đo để hiệu chỉnh lại ước đoán.



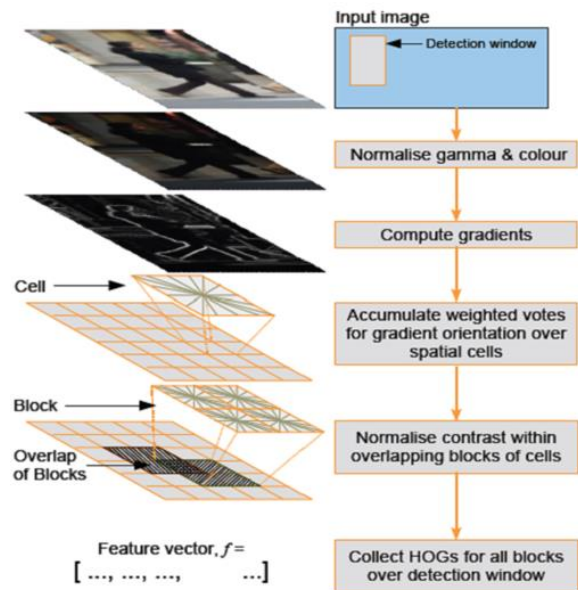
Hình 4: Thuật toán Kalman cổ điển

Trên thực tế, với những hệ thống có yếu tố phi tuyến thì chất lượng của bộ lọc Kalman tuyến tính là chưa tốt, vì thế bộ lọc Kalman mở rộng (EKF) được xem là một trong những cách hiệu quả để tăng cường chất lượng của quá trình ước lượng.

3.3. Nhận dạng hành động dùng 5-HOGs kết hợp với phân loại SVM

Histogram of gradient (HOG)[2] là đặc trưng được dùng nhiều trong lĩnh vực phát hiện đối tượng.

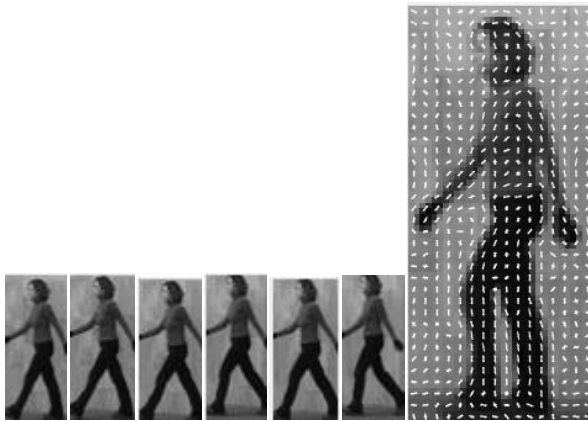
Ý tưởng chính trong đặc trưng HoG là hình dạng và trạng thái của vật có thể (1.2) đặc trưng bằng sự phân bố về gradient của cạnh. Đặc trưng này được phát triển dựa trên SIFT, đặc trưng HOG được tính trên cả một vùng. Do sự biến thiên màu sắc trong các vùng là khác nhau, kết quả là mỗi vùng sẽ cho ta một vector đặc trưng của nó. Vì vậy để có được đặc trưng của toàn bộ cửa sổ ta phải kết hợp nhiều vùng liên tiếp lại với nhau.



Hình 5: Thuật toán rút trích đặc trưng HOG [2]

Bản chất video là từ nhiều khung ảnh liên tiếp kết hợp lại với nhau theo một tốc độ nhất định. Để khảo sát hành động của người thông qua video ta tiến hành bám theo đối tượng. Dùng thông tin của các frame ảnh liên tiếp để xác định hành động của đối tượng bằng việc ghép nhiều đặc trưng HOG liên tiếp trong các khung ảnh quanh khu vực người khảo sát. Đây là một phương pháp đơn giản để phân loại chuyển động và cũng tồn tại các hạn chế

như độ chính xác chưa cao.



Hình 6: Đặc trưng HOG trong 5-frames ảnh và sắp xếp lại thành 1 đặc trưng về hành động. Sau khi trích được trung về hành động ta dùng bộ phân loại SVM để phân loại dữ liệu.

4. KẾT QUẢ THỰC HIỆN

Cơ sở dữ liệu:

- Bộ phát hiện được huấn luyện trên tập dữ liệu Caltech với khung hình có kích thước 50x21.

- Bộ phân loại hành động huấn luyện trên tập huấn luyện gồm 2090 mẫu các hành động “walk”, “run”, “side”, “skip”, “jump” trong tập dữ liệu Weizzman[17] kết hợp với mẫu tác giả thu thập thêm từ thực tế.

Phát hiện và bám theo đối tượng người:

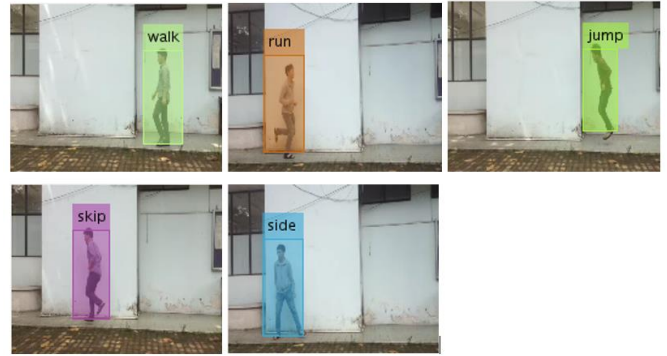
Hệ thống đã phát hiện và bám theo đối tượng người trong khung ảnh, với kích thước khung ảnh nhỏ, tốc độ xử lý khá nhanh có thể đáp ứng thời gian thực.



Hình 7: Kết quả sau khi phát hiện người và bám theo mục tiêu.

Hệ thống đã phân loại được 5 hành động khác nhau của đối tượng người: đi - chạy -

nhảy - lò cò – trượt ngang.



Hình 8: kết quả nhận dạng hành động

Kết quả nhận dạng các hành động khi đã phát hiện và bám mục tiêu. Test lại trên mẫu 934 gồm 314 mẫu “walk”, 110 mẫu “run” và 211 mẫu “side”, 170 mẫu, “skip”, 129 mẫu “jump”.

	Walk	Run	Side	Skip	Jump
Walk	0.97	0.00	0.00	0.01	0.02
Run	0.00	0.81	0.00	0.19	0.00
Side	0.05	0.01	0.93	0.00	0.00
Skip	0.00	0.15	0.00	0.85	0.00
Jump	0.00	0.02	0.00	0.05	0.93

Bảng 1: Kết quả nhận dạng hành động trên tập test

Chương trình thực thi trên Laptop có CPU Intel CoreI5-2520 – 2.50GHZ, Ram 12GHZ trên khung hình 176x144 và 320x240.

Kết quả nhận dạng hành động người là xâu chuỗi của ba giai đoạn: phát hiện người - bám theo mục tiêu và nhận dạng hành động người. Vì vậy, kết quả của từng giai đoạn ảnh hưởng trực tiếp đến kết quả chương trình.

5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

5.1. Kết luận

Dựa trên những kết quả của nghiên cứu trước, đề tài xây dựng hệ thống phát hiện và phân loại hành vi đối tượng.

- Hệ thống đã phát hiện và bám theo đối tượng là người trong khung hình.

- Hệ thống đã phân loại được các loại hành động khác nhau của đối tượng.

- Tốc độ phát hiện đối tượng và phân loại hành vi là tương đối trong khác khung ảnh có độ phân giải thấp.

Tuy nhiên, đề tài cũng còn tồn tại các hạn chế:

- Do bước phân loại hành động là tổng hợp của cả 3 giai đoạn phát hiện người, bám mục tiêu (tracking) và nhận dạng hành động nên kết quả của giai đoạn này là tài liệu để thực thi tiếp giai đoạn tiếp theo nên đòi hỏi các bước phát hiện người - bám mục tiêu (tracking) và phân loại đặc trưng hành động phải có độ chính xác đủ cao, và tốc độ xử lý đủ nhanh để việc thực thi có thể liền mạch.

- Độ chính xác là một thách thức lớn trong bài toán nhận dạng – phân loại hành động ở người.

5.2. Hướng phát triển của đề tài

Nghiên cứu, tìm hiểu, cải tiến các thuật toán khác để nâng cao độ chính xác trong phát hiện và bám theo đối tượng người.

- Phân loại nhiều loại hành động và phát triển tập mẫu.
- Phát hiện hành vi của đối tượng trong đêm khi kết hợp các loại camera và nhiều phương xử lý khác nhau.
- Kết hợp với các thiết bị khác để tăng hiệu quả trong việc hỗ trợ giám sát an ninh (báo động khi có hành vi bất thường,...)
- Xây dựng một hệ thống hoàn chỉnh từ phần cứng đến phần mềm để áp dụng ứng dụng giám sát an ninh thực tế.

TÀI LIỆU THAM KHẢO

- [1] Manoranjan Paul, Shah M E Haque and Subrata Chakraborty, “*Human detection in surveillance videos and its applications - a review*”, EURASIP, p.176, 2013.
- [2] N. Dalal; B. Triggs, “*Histograms of oriented gradients for human detection*”, Pages: 886 - 893 vol. 1, DOI: 10.1109/CVPR.2005.
- [3] Viola, Paul A. and Jones, Michael J. “*Rapid Object Detection using a Boosted Cascade of Simple Features*”, IEEE CVPR, 2001.
- [4] G.R. Bradski “*Real Time Face and Object Tracking as a Component of a Perceptual User Interface*”, Proceedings of the 4th IEEE Workshop on Applications of Computer Vision, 1998.
- [5] Shu, Guang, “*Human Detection, Tracking and Segmentation in Surveillance Video*”, *Electronic Theses and Dissertations*. Paper 4598.
<http://stars.library.ucf.edu/etd/4598> , Year 2014.
- [6] Fan Yang; Huchuan Lu; Ming-Hsuan Yang, “*Robust Superpixel Tracking*”, IEEE Transactions on Image Processing, Pages: 1639-1651, DOI: 10.1109/TIP.2014.2300823, Volume: 23, Issue: 4, Year: 2014.
- [7] I. Laptev and T. Lindeberg; “*Space-Time Interest Points*”, Nice, France, pp.I:432-439 in Proc. ICCV'03.
- [9] Lê Viết Tuấn, “*Nhận diện hành động người bằng cập nhật kết quả phân lớp*”, trang 10-14, bản tin khoa học & giáo dục 2014.
- [8] Christian Schuldt, Ivan Laptev and Barbara Caputo, “*Recognizing Human Action: A Local SVM Approach*”, ICPR'04 – Cambridge - UK, p.III:32—36, 2004.
- [10] Piotr Dollár; Ron Appel; Serge Belongie; Pietro Perona; “*Fast Feature Pyramids for Object Detection*”; Pages: 1532 – 1545. DOI: 10.1109/TPAMI.2014.2300479; Volume: 36, Issue: 8; Year: 2014.
- [11] Song Shao; Hong Liu; Xiangdong Wang; Yueliang Qian, “*Local Associated Features for Pedestrian Detection*”, Computer Vision - ACCV 2014 Workshops, Springer International Publishing, pp 513-526, 10.1007/978-3-319-16628-5_37, Year 2015.
- [12] Greg Welch and Gary Bishop, “*An Introduction to Kalman Filter*”, TR-95-041, Updated Monday, July, 24, 2006.
- [13] PETS, (2009). <http://ftp.pets.rdg.ac.uk>. Accessed 17 Nov 2013
- [14] M Blank, L Gorelick, E Shechtman, M Irani, R Basri, M Blank, L Gorelick, E Shechtman, M Irani, R Basri, “*Actions as space-time shapes*”, in Tenth IEEE International Conference on Computer Vision (ICCV '05) (IEEE, Piscataway, 2005), pp. 1395–1402
- [15] Carl Vondrick; Aditya Khosla; Tomasz Malisiewicz; Antonio Torralba, “*HOGgles: Visualizing Object Detection Features*”, 2013 IEEE International Conference on Computer Vision, Pages: 1 - 8, DOI: 10.1109/ICCV.2013.8, Year: 2013.
- [16] Pedro Felzenszwalb; David McAllester; Deva Ramanan, “*A discriminatively trained, multiscale, deformable part model*”, Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, Pages: 1 - 8, DOI: 10.1109/CVPR.2008.4587597, Year: 2008
- [17] M Blank, L Gorelick, E Shechtman, M Irani, R Basri, M Blank, L Gorelick, E Shechtman, M Irani, R Basri, “*Actions as space-time shapes*”, in Tenth IEEE International Conference on Computer Vision (ICCV '05) (IEEE, Piscataway, 2005), pp. 1395–1402, Year 2005.

Thông tin liên hệ tác giả chính (người chịu trách nhiệm bài viết):

Họ tên: Hồ Quốc Thiên

Đơn vị: Trường Đại học sư phạm kỹ thuật TP.HCM

Điện thoại: 0949 009 666

Email: hoquocthiendl@gmail.com

BÀI BÁO KHOA HỌC

THỰC HIỆN CÔNG BỐ THEO QUY CHẾ ĐÀO TẠO THẠC SĨ

Bài báo khoa học của học viên

có xác nhận và đề xuất cho đăng của Giảng viên hướng dẫn



Bản tiếng Việt ©, TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HỒ CHÍ MINH và TÁC GIẢ

Bản quyền tác phẩm đã được bảo hộ bởi Luật xuất bản và Luật Sở hữu trí tuệ Việt Nam. Nghiêm cấm mọi hình thức xuất bản, sao chụp, phát tán nội dung khi chưa có sự đồng ý của tác giả và Trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh.

ĐỂ CÓ BÀI BÁO KHOA HỌC TỐT, CẦN CHUNG TAY BẢO VỆ TÁC QUYỀN!

Thực hiện theo MTCL & KHTHMTCL Năm học 2017-2018 của Thư viện Trường Đại học Sư phạm Kỹ thuật Tp. Hồ Chí Minh.