**ĐẠI HỌC QUỐC GIA TP HCM**

TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN



**TRỊNH NAM VIỆT**

**ƯỚC LƯỢNG MÂT ĐỘ ĐÁM ĐÔNG**

**TRONG VIDEO GIÁM SÁT**

**LUẬN VĂN THẠC SĨ CÔNG NGHỆ THÔNG TIN**

Chuyên ngành: KHOA HỌC MÁY TÍNH

Mã số: 60.48.01.01

**TP HỒ CHÍ MINH – NĂM 2017**

**ĐẠI HỌC QUỐC GIA TP HCM**

TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN





**TRỊNH NAM VIỆT**

**ƯỚC LƯỢNG MẬT ĐỘ ĐÁM ĐÔNG**

**TRONG VIDEO GIÁM SÁT**

**LUẬN VĂN THẠC SĨ CÔNG NGHỆ THÔNG TIN**

Chuyên ngành: KHOA HỌC MÁY TÍNH

Mã số: 60.48.01.01

**NGƯỜI HƯỚNG DẪN KHOA HỌC:**

**TS. NGÔ ĐỨC THÀNH**

**TP HỒ CHÍ MINH – NĂM 2017**

**LỜI CAM ĐOAN**

Tôi xin cam đoan rằng luận văn thạc sĩ này là công trình nghiên cứu của riêng tôi. Các số liệu, bảng biểu có nguồn gốc, dẫn chứng rõ ràng. Kết quả thu thập được từ quá trình nghiên cứu là trung thực, không có sự sao chép từ đề tài, luận văn hay công trình nghiên cứu khoa học của tác giả nào khác.

*TP.HCM, tháng 11 năm 2017*

**Tác giả luận văn**

**TRỊNH NAM VIỆT**

**LỜI CẢM ƠN**

Trước tiên tôi xin gửi lòng biết ơn chân thành và sâu sắc đến TS Ngô Đức Thành, người đã giúp tôi định hướng đề tài và tận tình chỉ dẫn để tôi có thể hoàn thành được luận văn này.

Tôi xin cảm ơn các Thầy Cô của trường Đại học Công Nghệ Thông Tin đã tận tâm giảng dạy cho tôi những kiến thức quí giá trong các năm học cao học.

Tôi xin cảm ơn phòng quản lí và đào tạo sau đại học trường Đại học Công Nghệ Thông Tin đã bảo ban, nhắc nhở và tạo điều kiện thuận lợi trong suốt quá trình học tập.

Tôi xin cảm ơn gia đình, bạn bè, đồng nghiệp đã luôn ủng hộ, động viên và giúp đỡ tôi có thể hoàn thành khóa học và luận văn này.

*TP.HCM, tháng 11 năm 2017*

Trịnh Nam Việt

MỞ ĐẦU

Hiện nay, với sự phát triển vũ bão của khoa học công nghệ, các hệ thống camera giám sát ngày càng phổ biến với chất lượng hình ảnh ngày càng nâng cao và chi phí triển khai ở mức chấp nhận được. Các hệ thống camera giám sát thường được dùng để theo dõi và đánh giá tình hình an ninh trật tự ở các địa điểm công cộng (công viên, trường học, quảng trường, nhà ga, sân bay, …), các cơ quan trọng yếu (tòa nhà, cơ quan công quyền, khu vực an ninh/ quốc phòng). Đối tượng của các hệ thống camera giám sát này chủ yếu là con người và phương tiện giao thông.

Ngoài ra, thế giới hiện nay diễn biến phức tạp, tình trạng khủng bố giết người hang loạt vào đám đông xảy ra khắp mọi nơi. Tình trạng kẹt xe trầm trọng gây thiệt hại nghiêm trọng cho nhiều nền kinh tế, riêng thành phố Hồ Chí Minh bị thiệt hại khoảng 25.000 tỉ đồng mỗi năm do tình trạng kẹt xe gây ra. Do đó với hai đối tượng giám sát chủ yếu là con người và phương tiện thì hành vi đám đông cần được ưu tiên giám sát, để từ đó có thể phần tính đánh giá đúng tình hình an ninh trật tự, lưu lượng giao thông trong phạm vi giám sát của camera.

Với công nghệ hiện tại, việc giám sát đám đông qua camera vẫn chủ yếu dựa vào con người, tuy nhiên tại một thời điểm thì một người chỉ có thể tập trung giám sát được số lượng camera giới hạn. Vì vậy việc xây dựng xây dựng hệ thống có thể phát hiện một cách chính xác mật độ đám đông của người và phương tiện trong hình ảnh hoặc video để từ đó có thể điều tiết, quả lí và đảm bảo an ninh công cộng là một bài toán khó và rất cấp thiết.

Trong pham vị luận văn này, tác giả sẽ tìm hiểu những phương pháp tốt để thực hiện việc ước lượng mật độ đám đông, từ đó cài đặt lại và so sánh với kết quả của những phương pháp khác. Từ đó có nền tảng kiến thức để tìm các cải tiến các phương pháp và xây dựng các ứng dụng dựa trên kết quả đã đạt được trong tương lai.

DANH MỤC KÍ HIỆU, TỪ VIẾT TẮT

DANH MỤC HÌNH

DANH MỤC BẢNG BIỂU

Chương 1. TỔNG QUAN

1.1 Lí do chọn đề tài

Mong muốn sống trong hòa bình, hanh phúc và thịnh vượng là khát vọng chính đáng của loài người. Tuy nhiên, thế giới hiện nay có nhiều bất ổn với diễn biến phức tạp, đặc biệt là chủ nghĩa khủng bố lây lan như dịch bệnh, gieo rắc đau thương cho nhân loại. Dân số bùng nổ dẫn đến hạ tầng giao thông nhiều quốc gia không đáp ứng được nhu cầu tham gia giao thông của số lượng lớn người và phương tiện, dẫn đến thiệt hại kinh tế nặng nề cho nền kinh tế của các quốc gia này.

Những vấn đề trên nếu không được quản lí, giải quyết hiệu quả sẽ gây ảnh hưởng sống còn đến loài người. Do đó các quốc gia khắp nơi trên thế giới đã cung cấp rất nhiều về nhân lực cũng như tài lực để phòng ngừa cũng như ngăn chặn ảnh hưởng của chúng. Tuy nhiên, việc quản lí phòng ngừa chủ yếu do con người thực hiện mà khả năng của con người có hạn, chỉ có thể hạn chế được một phần. Nên việc có được một hệ thống tự động ước lượng được mật độ đám đông của người và phương tiện là cấp thiết.

Từ năm 2012, khi mô hình CNN AlexNet giành được chiến thắng vang dội tại cuộc thi ImageNet Large Scale Visual Recognition Challenge [ref] thì thì thuật ngữ Deep Learning đã trở nên phổ biến, những bài báo về Deep Learning chiếm tỉ lệ lớn trong các công trình được được công bố. Từ đó Deep Learning đã được xem như là tối ưu trong các mô hình dự đoán máy học.

Ngoài ra với sự phát triển như nhanh chóng của khoa học công nghệ, các thiết bị ghi hình và lưu trữ ngày càng nâng cao chất lượng với chi phí giảm xuống. Vì vậy việc triển khai các hệ thống camera giám sát có thể được triển khai nhanh chóng với chi phí thấp.

Từ những yếu tố thuận lợi trên, phát hiện và phân tích đám đông đã thu hút công đồng nghiên cứu trong thời gian gần đây [12]. Hướng nghiên cứu này tập trung giải quyết nhiều bài toán con như: phân tích tính ổn định và nhất quán trong chuyển động của đám đông [1-3]; phát hiện vị trí đám đông [7]; nhận dạng và phân tích hoạt động, hành vi đám đông [4-5,8-9,11,13]; phát hiện sự kiện bất thường [6]; đếm số lượng đối tượng và ước lượng mật độ [14,15-17].... Trong đó, ước lượng mật độ là một trong những bước quan trọng để làm cơ sở cho các chức năng phân tích hành vi, phát hiện sự kiện bất thường. Đây cũng là chủ đề nghiên cứu mà học viên tập trung thực hiện trong luận văn này.

1.2 Mục tiêu của đề tài

Tìm hiểu các phương pháp tiên tiến đã được công bố trể các tập dữ liệu chuẩn.

Cài đặt lại và thử nghiệm phương pháp CCNN, Hydra CNN [16], từ kết quả thực nghiệm so sánh với kết quả của các phương pháp tiên tiến khác.

Từ kết quả đạt được được tìm cách cải tiến nâng cao hiệu suất.

Xây dựng ứng dụng thử nghiệm

1.3 Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu là các các phương pháp ước lượng mật độ đám đông (con người hoặc phương tiện) trong các đoạn video giám sát.

Phạm vi nghiên cứu là tiến hành thực nghiệm và đánh giá trên các tập dữ liệu về người là UCF\_CC\_50, UCSD; và tập dữ liệu về phương tiện là TRANCOS.

1.4 Phương pháp nghiên cứu

Cài đặt lại các phương pháp CCNN, Hydra CNN.

Thực nghiệm và đánh giá trên các bộ dữ liệu UCF\_CC\_50, UCSD, TRANCOS.

Từ kết quả quả thực nghiệm, so sánh với kết quả của các phương pháp tiên tiến hiện có [14, 15, 16, 17]

Phân tích tìm cách cải tiến nâng cao hiệu suất của các phương pháp đã cài đặt lại

1.5 Ý nghĩa của đề tài

Tìm hiểu và xây dựng mô hình ước lượng mật độ đám đông một cách chính xác là tiền đề để tác giả tiếp tục nghiên cứu xây dựng các ứng dụng hỗ trợ giám sát trong tương lai

Chương 2. CƠ SỞ LÍ THUYẾT

Ở chương này trình bày những định nghĩa, khái niệm và các phương pháp ước lượng; khái niệm về camera giám sát; khái niệm về đám đông và phương pháp trực quan hóa mật độ đám đông từ đám đông đầu vào; lý thuyết về mạng CNN và các mô hình sử dụng mạng CNN để ước lượng mà luận văn áp dụng để giải quyết bài toán ước lượng mật độ đám đông.

2.1 Tổng quan về ước lượng

2.1.1 Khái niệm về ước lượng

Trong thống kê, một ước lượng là một giá trị được tính toán từ một mẫu thử (échantillon) và người ta hy vọng đó là giá trị tiêu biểu cho giá trị cần xác định trong tổng thể (population). Người ta luôn tìm một ước lượng sao cho đó là ước lượng "không chệch" (unbiased), hội tụ (converge), hiệu quả (efficient) và vững (robust).

2.1.2 Một số thuộc tính trong ước lượng

* **Giá trị kì vọng**

Giá trị kỳ vọng, giá trị mong đợi (hoặc kỳ vọng toán học), hoặc trung bình của một biến ngẫu nhiên - viết tắt là µ - là trung bình có trọng số của tất cả giá trị có thể của biến đó. Hay là được tính bằng tổng các tích giữa xác suất xảy ra của mỗi giá trị có thể của biến với giá trị đó. Định nghĩa theo ngôn ngữ toán học là như sau:

Cho X ngẫu nhiên với một số hữu hạn các kết quả hữu hạn x1,x2, …, xn xảy ra với các xác suất p1, p2, …, pn tương ứng. Kì vọng của X được định nghĩa là:

**E[X] = x1p1 + x2p2 + … + xnpn**

Từ đó đó nó biểu diễn giá trị trung bình mà người ta mong đợi thắng cược khi đặt cược nhiều lần với tỉ lệ thắng là như nhau. Khi đó, một trò chơi được coi là công bằng khi giá trị kì vọng = 0

VD: 1 vòng quay rô-lét có 38 ô, có khả năng quay trúng là như nhau, khi trúng vào ô đặt sẽ thắng số tiền gấp 35 lần cùng với số tiền đặt cược được hoàn lại là 36 lần tiền cược. Khi đó ta xét giá trị kì vọng cho 38 ô với giá trị đặt cược là 1 đô la sẽ được như sau:

Như vậy kết quả sẽ là bạn sẽ mất trung bình khoảng hơn 5 xu một chút cho mỗi 1 đô la đặt cược.

2.1.3 Đánh giá ước lượng

Một ước lượng là một giá trị x (x nhỏ) được tính toán trên một mẫu được lấy một cách ngẫu nhiên, do đó giá trị của x là một biến ngẫu nhiên với kì vọng E(x) và phương sai V(x). Nghĩa là giá trị x có thể dao động tùy theo mẫu thử, nó có ít cơ hội để có thể bằng đúng chính xác giá trị X (X lớn) mà nó đang ước lượng. Mục đích ở đây là ta muốn có thể kiểm soát sự sai lệch giá trị x và giá trị X.

Một biến ngẫu nhiên luôn dao động xung quanh giá trị kì vọng của nó. Ta muốn là kì vọng của x phải bằng X. Khi đó ta nói ước lượng là không chệch (unbiased).

Ta cũng muốn là khi mẫu thử càng rộng, thì sai lệch giữa x và X càng nhỏ. Khi đó ta nói ước lượng là hội tụ. Định nghĩa theo ngôn ngữ toán học là như sau:

hội tụ nếu với mọi số thực **ϵ** dương

(xác suất để sai lệch với giá trị thực cần ước lượng lớn hơn **ϵ** tiến về 0 khi kích cỡ của mẫu thử càng lớn)

Biến ngẫu nhiên dao động quanh giá trị kì vọng của nó. Nếu phương sai V(x) càng bé, thì sự dao động càng yếu. Vì vậy ta muốn phương sai của ước lượng là nhỏ nhất có thể. Khi đó ta nói ước lượng là hiệu quả (eficient).

Cuối cùng, trong quá trình điều tra, có thể xuất hiện một giá trị "bất thường" (ví dụ có trẻ 10 tuổi nhưng cao 1,80 m). Ta muốn giá trị bất thường này không ảnh hưởng quá nhiều đến giá trị ước lượng. Khi đó ta nói ước lượng là vững (robust).