Người ta thường nghĩ cảm xúc là thứ xa xỉ, thứ không cần thiết cho hoạt động cơ bản và khó mã hóa đối với một chương trình máy tính. Vậy tại sao ta lại cần bận tâm đến việc trao cho máy móc khả năng cảm xúc? Trong thời gian gần đây, rất nhiều các phát hiện, từ khoa học thần kinh, tâm lý học đến khoa học nhận thức, cho thấy rằng cảm xúc đóng những vai trò quan trọng đến đáng ngạc nhiên trong việc thực hiện hành vi một cách đúng đắn và hợp lý. Chúng ta đều biết rằng quá nhiều cảm xúc sẽ gây ảnh hưởng không tốt cho việc đưa ra quyết định hợp lý trong nhiều trường hợp. Tuy nhiên các nghiên cứu khoa học thần kinh trên những bệnh nhân trong tình trạng cơ bản bị mất cảm xúc đã cho thấy rằng việc tính đúng đắn khi đưa ra quyết định trong những hoạt động đòi hỏi trí thông minh của những bệnh nhân này bị suy giảm mạnh, từ đó đưa ra nhận định rằng quá ít cảm xúc có thể làm giảm khả năng tư duy và thực hiện hành vi hợp lý. Có thể thấy rằng, cảm xúc tương tác với suy nghĩ theo những cách không rõ ràng nhưng lại đóng vai trò quan trọng đối với hoạt động đòi hỏi trí thông minh. Các vùng não xử lý cảm xúc cũng được phát hiện thực hiện các quá trình nhận dạng khuôn mẫu trước khi chuyển tiếp các tín hiệu đến đến vỏ não. Ví dụ, một con chuột có thể có phản xạ sợ một âm thanh ngay cả khi vỏ não thính giác của nó đã bị loại bỏ; quá trình xử lý theo định hướng cảm xúc tương tự cũng được cho là diễn ra trong hoạt động nghe nhìn của con người.

Các nhà khoa học đã thu thập được bằng chứng cho thấy khả năng cảm xúc là một thành phần cơ bản của trí thông minh, đặc biệt là đối với khả năng học tập và thích ứng. Với việc triển khai ngày càng nhiều các hệ thống máy tính thích ứng, điển hình là các hệ thống phần mềm thu thập video để học hỏi từ người dùng, khả năng cảm nhận và phản hồi phù hợp với hành động cũng như cảm xúc người dùng ngày càng quan trọng. Trí tuệ cảm xúc bao gồm khả năng nhận biết, thể hiện và có cảm xúc, cùng với đó khả năng điều chỉnh những cảm xúc này. Cần khai thác điều này cho các mục đích xây dựng và xử lý khéo léo đối với cảm xúc của người khác. Trí tuệ cảm xúc cũng đã được cho là một yếu tố dự đoán tốt hơn chỉ số IQ trong việc đo lường các khía cạnh của sự thành công trong cuộc sống.

Máy móc có thể sẽ không bao giờ cần đến tất cả các khả năng cảm xúc mà con người có; tuy nhiên, có bằng chứng cho thấy máy móc sẽ yêu cầu một số cảm xúc thiết yếu để tỏ ra thông minh khi tương tác với con người. Một giả thuyết liên quan của Reeves và Nass tại đại học Stanford cho rằng: Tương tác giữa người với máy tính vốn dựa trên tự nhiên cũng như quá trình giao tiếp xã hội, nó tuân theo những điều cơ bản nhất trong tương tác giữa người với người. Ví dụ, nếu một thiết bị công nghệ giao tiếp với bạn nhưng không bao giờ lắng nghe bạn thì việc nó gây phiên phức cho bạn là điều khó tránh khỏi, tương tự như việc một người nói chuyện với bạn nhưng không bao giờ lắng nghe đến cảm xúc của bạn trong khi giao tiếp. Nass và Reeves đã tiến hành hàng chục thí nghiệm về sự tương tác giữa người với người cổ điển, mô phỏng lại trong máy tính và nhận thấy rằng các kết quả cơ bản trong tương tác giữa người với người vẫn được giữ nguyên.

Nhận biết những phản hồi liên quan đến cảm xúc là điều quan trọng đối với sự tương tác giữa con người và máy tính thông minh. Hãy xem xét một thuật toán học máy phải quyết định thời điểm nào nên ngắt lời người dùng. Con người quyết định điều này bằng cách đánh giá phản ứng của đối phương khi bị ngắt lời trong các tình huống khác nhau: việc này gây ra phản ứng trung lập, tiêu cực hay tích cực? Nếu không quan tâm đến phản ứng của đối phương, người đó có thể bị coi là thiếu tôn trọng, một kẻ khó chịu hay thậm chí là thiếu thông minh. Từ đó ta có thể dự đoán một phản ứng tương tự dánh cho máy tính dựa trên biểu hiện tích cực hay tiêu cực trong phản hồi của người dùng giúp máy tính có thể tỏ ra thông minh hơn.

Mặc dù không phải tất cả các máy tính đều cần các khả năng cảm xúc, nhưng đối với những máy tính có tương tác và cần thích ứng với con người trong thời gian thực có thể được coi là thông minh hơn nếu có khả năng này. Phần còn lại của luận văn này sẽ tập trung vào việc cung cấp cho máy móc một trong những khả năng quan trọng của trí tuệ cảm xúc , đó là khả năng nhận biết cảm xúc của người dùng.

1.1 Nhận biết cảm xúc con người

Trẻ sơ sinh cho thấy khả năng nhận biết các biểu hiện tình cảm như tán thành và không tán thành rất lâu trước khi chúng hiểu ngôn ngữ. Nhận biết biểu cảm được cho là đóng một vai trò quan trọng trong việc học tập và phát triển. Nó có thể là một phần quan trọng trong việc đánh giá sự khác biệt giữa sự phát triển bình thường của trẻ em so với trẻ tự kỷ, những người thường bị suy giảm khả năng nhận biết biểu cảm. Ví dụ, thay vì chăm chú vào bài phát biểu của cha mẹ với sự suy diễn quá mức, trẻ tự kỷ có thể điều chỉnh âm thanh chúng cho là không quan trọng do thiếu sự định hướng của các tín hiệu cảm xúc.

Cảm xúc điều chỉnh hầu hết các hành vi giao tiếp của con người: lựa chọn từ ngữ, giọng nói, nét mặt, hành vi cử chỉ, tư thế, nhiệt độ da và độ nhờn, hô hấp, căng cơ, v.v. Cảm xúc có thể góp phần thay đổi đáng kể nội dung thông điệp: đôi khi điều quan trọng nhất không phải là nội dung câu nói mà là cách thể hiện câu nói. Khuôn mặt có xu hướng là hình thức giao tiếp cảm xúc dễ nhìn thấy nhất, nhưng chúng cũng dễ dàng kiểm soát nhất để phản ứng phù hợp với các tình huống giao tiếp khác nhau nếu so sánh với giọng nói và các phương thức biểu đạt khác. Nhận biết biểu cảm có độ chính xác cao nhất khi nó kết hợp với: 1) Nhận được nhiều các loại tín hiệu khác nhau 2) Nắm bắt được ngữ cảnh, tình huống, mục tiêu và sở thích của đối phương. Từ đó, việc kết hợp giữa nhận dạng mẫu cấp thấp, lập luận cao cấp cùng với xử lý ngôn ngữ tự nhiên sẽ đưa ra đánh giá về biểu cảm, cảm xúc một cách chuẩn xác nhất.

Máy tính sẽ phải nhận biết trạng thái cảm xúc của con người tốt đến mức nào? Cần lưu ý rằng không ai có thể nhận ra cảm xúc sâu kín nhất của đối phương một cách hoàn hảo, thậm chí đôi khi con người còn không thể nhận ra cảm xúc của chính họ. Không có phương thức giao tiếp đã được biết đến nào được coi là hoàn hảo. Một số khía cạnh của cảm xúc bên trong vẫn được giữ kín, đặc biệt nếu đó là chủ đích hoặc nếu chúng được ngụy trang một cách kĩ càng. Những gì có thể quan sát cũng như đánh giá luôn đi kèm với một mức độ thiếu chắc chắn nhất định. Tuy nhiên, người ta vẫn có thể nhận ra cảm xúc của nhau ở mức độ đủ để trao đổi thông tin và phản hồi một cách phù hợp. Mục đích của chúng tôi là cung cấp cho máy tính khả năng nhận diện tương tự hoặc gần bằng khả năng mà con người có.

1.2 Tầm quan trọng của nhận biết Cảm xúc sinh lý học

Khi thiết lập các giao diện máy thông minh, tại sao không tập trung vào giao tiếp bằng khuôn mặt và giọng nói — đây rõ ràng là những phương thức mà con người thường sử dụng nhất? Có một vài trường hợp khả năng như vậy sẽ mang lại hiệu quả hơn. Tuy nhiên, sẽ là một sai lầm khi nghĩ về sinh lý như một cái gì đó mà con người không thể nhận ra một cách tự nhiên. Một cái bắt tay từ có thể cho bạn cảm nhận được độ thiện chí của nó; một người đứng cạnh có thể cảm thấy nhịp tim của bạn; sinh viên có thể nghe thấy những thay đổi trong tiếng thở của giáo sư, cho thấy căng thẳng; hay những biểu hiện cảm xúc có thể được thể hiện qua sức căng của cơ mặt… Cho thấy rằng, con người có thể nắm bắt được cảm xúc đối phương thông qua các biểu hiện sinh lý học.

Nhận biết các dạng biểu hiện sinh lý học của cảm xúc có ứng dụng quan trọng trong y học, giải trí và cả trong tương tác giữa con người với máy tính. Các trạng thái như suy sụp, lo lắng và tức giận mãn tính đã được chứng minh là cản trở hoạt động của hệ thống miễn dịch, khiến con người dễ bị nhiễm bệnh hơn và làm chậm quá trình hồi phục sau chữa trị hoặc phẫu thuật. Nhận biết các biểu hiện sinh lý có thể hỗ trợ trong việc đánh giá và định lượng sự căng thẳng, tức giận cũng như các cảm xúc khác gây ảnh hưởng xấu đến sức khỏe. Một số biểu hiện sinh lý của con người cho thấy phản ứng đặc trưng với âm nhạc và các hình thức giải trí khác, ví dụ: độ nhạy cảm của da có xu hướng tăng cao khi một bản nhạc khiến bạn hưng phấn và giảm xuống khi nó khiến bạn trấn tĩnh lại. Nguyên tắc này đã được sử dụng trong một thiết bị đeo có tên "DJ cảm xúc" nó cho phép lựa chọn các bài hát phù hợp với từng người dùng. Những thay đổi trong tín hiệu sinh lý cũng có thể được theo dõi để tìm các dấu hiệu căng thẳng phát sinh trong khi người dùng tương tác với thiết bị, giúp phát hiện liệu sản phẩm có gây ra sự khó chịu cho người dùng hay không trong khi không cần phải làm gián đoạn trải nghiệm người dùng. Đây là một lĩnh vực mới dành cho nghiên cứu nhận dạng mẫu: phát hiện khi nào sản phẩm gây ra sự khó chịu cho người dùng, từ đó giúp các nhà phát triển thay đổi thiết kế cũng như cải tiến.

Cảm biến sinh lý đôi khi được coi là xâm lấn vì nó liên quan đến tiếp xúc vật lý với con người. Tuy nhiên, công nghệ ngày càng được cải tiến với các điện cực vải hay cao su dẫn điện có thể đeo hoặc giặt, chúng có thể được tích hợp ngay trong quần áo và phụ kiện hằng ngày. Thêm vào đó còn có những hình thức hoạt động mới của cảm biến sinh lý khi không yêu cầu tiếp xúc. Trong một số trường hợp, cảm biến sinh lý còn được cho là ít xâm lấn hơn các lựa chọn khác, chẳng hạn như so với quay video. Sinh viên tham gia đào tạo từ xa có thể cung cấp thông tin cho giảng viên về các trạng thái của họ nhưng giảng viên không biết danh tính của họ. Vì vậy sẽ ít có sự phản đối hơn do thông tin có thể được truyền đi một cách ẩn danh.

Các thể máy tính đeo tay mới tạo điều kiện cho các loại cảm biến nhiều hơn so với máy tính truyền thống. Chúng thường xuyên có khả năng tiếp xúc tự nhiên với bề mặt của da; tuy nhiên lại gặp vấn đề về việc ghi hình khuôn mặt người dùng. Trong các hệ thống đeo được, cảm biến sinh lý sẽ được thiết kế hết sức nhỏ gọn đến mức khó nhận biết bằng mắt thường. Đây là một điểm ưu việt so với việc sử dụng video.

1.3 Một sô nghiên cứu liên quan

Vấn đề nhận dạng biểu cảm là một vấn đề khó khi bạn nhìn vào một vài tiêu chuẩn tồn tại. Nói chung, con người có thể nhận ra một biểu hiện cảm xúc với độ chính xác khoảng 60%, với thử nghiệm có 6 lựa chọn. Các thuật toán máy tính có thể đạt tới hoặc thậm chí có kết quả tốt hơn tốt. Cũng cần lưu ý rằng tính năng nhận dạng giọng nói của máy tính hoạt động với độ chính xác khoảng tới 90% đối với giọng nói bình thường có xu hướng giảm độ chính xác xuống 50-60% đối với lời nói chứa đựng cảm xúc. Bởi vậy cải thiện khả năng xử lý cảm xúc trong lời nói là rất quan trọng trong lĩnh vực này.

Nhận dạng nét mặt dễ dàng hơn đối với mọi người, độ chính xác lên 70-98% trên sáu loại biểu hiện trên khuôn mặt của các diễn viên và tỷ lệ máy tính thu được độ chính xác từ 93-98% khi nhận dạng 5-7 kiểu biểu hiện cảm xúc trên các nhóm 8-32 người. Các nghiên cứu khác không tập trung quá nhiều vào việc nhận biết cảm xúc mà ưu tiên vào việc nhận biết các biểu hiện cụ thể trên khuôn mặt — các chuyển động cơ bản được mô tả trong hệ thống mã hóa hành động trên khuôn mặt của Paul Ekrnan — và có thể được sử dụng kết hợp để mô tả tất cả các biểu hiện trên khuôn mặt. Các bộ nhận dạng đã được tạo ra cho một số ít các hành động trên khuôn mặt và các bộ nhận dạng tự động đã được chứng minh là hoạt động tương đương với những người được đào tạo về nhận dạng các biểu hiện trên khuôn mặt. Những biểu hiện trên khuôn mặt về cơ bản là âm vị \* id, có thể được kết hợp lại để tạo thành biểu cảm trên khuôn mặt. Những nỗ lực gần đây còn chỉ ra rằng việc kết hợp các tín hiệu âm thanh và video để ghi nhận cảm xúc có thể cho kết quả cải thiện.

Mặc dù sự tiến bộ trong nhận dạng khuôn mặt, giọng nói cũng như kết hợp cả hai là đầy hứa hẹn, song tất cả các kết quả ở trên đều dựa trên dữ liệu được phân tích trước của một tập hợp các biểu cảm đôi khi được phóng đại quá mức hoặc trên một nhóm nhỏ các hành động khuôn mặt đơn lẻ được gán nhãn sẵn. Sự phát triển trong nhận dạng biểu cảm có chiều hướng tương tự như nhận dạng giọng nói vài thập kỷ trước khi máy tính có thể phân loại các chữ số được khớp nối cẩn thận được nói chậm với các khoảng dừng ở giữa, nhưng không thể phát hiện chính xác các chữ số này các chữ số theo nhiều cách mà chúng được nói trong các cuộc trò chuyện liên tục.

Nghiên cứu nhận dạng cảm xúc cũng phức tạp hơn vì hiểu được cảm xúc là rất khó. Sau hơn một thế kỷ nghiên cứu, các nhà nghiên cứu về cảm xúc vẫn chưa thống nhất về cảm xúc là gì và chúng được truyền đạt như thế nào. Một trong những câu hỏi lớn trong nghiên cứu cảm xúc là liệu các biểu hiện sinh lý khác biệt có tương ứng với mỗi cảm xúc hay không. Các chuyển động cơ sinh lý bao gồm biểu hiện trên khuôn mặt mà người ngoài nhìn thấy có thể không phải lúc nào cũng tương ứng với trạng thái cảm xúc cơ bản thực sự. Cảm xúc bao gồm nhiều thứ hơn là biểu hiện bên ngoài của nó; nó cũng bao gồm những cảm xúc và suy nghĩ bên trong, thêm vào đó là việc tồn tại các quá trình bên trong khác mà người có cảm xúc đó có thể không nhận thức được.

Mối quan hệ giữa cảm giác bên trong cơ thể và biểu hiện bên ngoài có thể quan sát được vẫn còn là một lĩnh vực nghiên cứu mở, với rất nhiều tranh cãi. Trong lịch sử, James là người đề xướng chính cho nhận định cảm xúc như một trải nghiệm về những thay đổi của cơ thể, chẳng hạn như tim bạn đập mạnh hoặc tay bạn đổ mồ hôi. Quan điểm này đã bị thách thức bởi Cannon và một lần nữa bởi Schachter và Singer, những người cho rằng trải nghiệm về những thay đổi sinh lý không đủ để phân biệt cảm xúc. Các thí nghiệm của Schachter và Singer cho thấy rằng nếu một trạng thái kích thích cơ thể được tạo ra, thì các đối tượng có thể được đưa vào hai tâm trạng riêng biệt chỉ đơn giản bằng cách đặt vào hai tình huống khác nhau. Họ lập luận rằng các phản ứng sinh lý như lòng bàn tay đổ mồ hôi và nhịp tim đập nhanh thông báo cho não bộ của chúng ta rằng chúng ta đang bị kích thích và sau đó bộ não phải đánh giá tình huống của trước khi nó có thể đánh trạng thái đó bằng một cảm xúc cụ thể sợ hãi hoặc yêu thương.

Kể từ nghiên cứu của Schachter và Singer, đã có một cuộc tranh luận về việc liệu cảm xúc có đi kèm với những thay đổi sinh lý cụ thể không chỉ đơn giản là mức độ kích thích. Ekman và Winton đã cung cấp một số phát hiện đầu tiên cho thấy sự khác biệt đáng kể trong các tín hiệu của hệ thần kinh tự chủ theo một số lượng nhỏ các loại hoặc chiều cảm xúc, nhưng không có sự khám phá về sự phân loại tự động. Fridlund và Izard dường như là những người đầu tiên áp dụng nhận dạng khuôn để phân loại cảm xúc từ các đặc điểm sinh lý, đạt được tỷ lệ chính xác 33-51% trên phân loại phụ thuộc vào chủ thể của bốn biểu hiện trên khuôn mặt (vui, buồn, tức giận, sợ hãi) cho bốn tín hiệu điện cơ đồ trên khuôn mặt. Mặc dù có khá nhiều kết quả đã được công bố nhằm mục đích tìm kiếm các tương quan sinh lý khi kiểm tra các nhóm cảm xúc nhỏ (từ 2-7 cảm xúc), hầu hết đều tập trung vào Nest hoặc phân tích so sánh phương sai, kết hợp dữ liệu trên nhiều đối tượng, trong đó mỗi đối tượng được đo trong một khoảng thời gian tương đối nhỏ (giây hoặc phút). Tương đối ít nghiên cứu bao gồm các trạng thái kiểm soát trung lập trong đó đối tượng thư giãn và trôi qua thời gian không cảm thấy có cảm xúc cụ thể nào và không có thông tin về việc liệu có nghiên cứu nào thu thập dữ liệu từ một người liên tục, trong nhiều tuần, nơi các nguồn nhiễu khác nhau nhập vào dữ liệu.

Công trình trong bài báo này là mới lạ trong việc cố gắng phân loại các mẫu biểu hiện sinh lý cho một nhóm tám cảm xúc (bao hàm trung tính), bằng cách áp dụng các kỹ thuật phân loại theo mẫu ngoài kỹ thuật phân biệt đơn giản cho vấn đề và bằng cách tập trung vào cảm xúc "cảm nhận được" của một đối tượng duy nhất được thu thập qua các phiên kéo dài nhiều tuần. Kết quả thu được cũng không phụ thuộc vào các cuộc tranh luận tâm lý về tính phổ quát của các loại cảm xúc thay vào đó tập trung vào các loại cảm xúc do người dùng xác định.

Những đóng góp của bài báo này bao gồm một phương thức mới để phân tích mô hình của các trạng thái cảm xúc từ sinh lý học, cũng như tìm ra tỷ lệ phân loại đáng kể từ các mô hình sinh lý tương ứng với tám trạng thái cảm xúc được quan sát và đánh giá đến từ một đối tượng trong nhiều tuần. Kết quả cũng cho thấy sự phân biệt đáng kể giữa cả hai chiều kích cảm xúc được mô tả phổ biến nhất: giá trị và kích thích. Chúng tôi chỉ ra rằng sự thay đổi của từ ngày này qua ngày khác trong các tín hiệu sinh lý là rất lớn, ngay cả khi cùng một cảm xúc được thể hiện và hiệu ứng này làm suy yếu độ chính xác của nhận dạng nếu nó không được xử lý thích hợp.