MỤC LỤC

[DANH MỤC CÁC HÌNH VẼ 3](#_Toc168515994)

[DANH MỤC CÁC BẢNG BIỂU 3](#_Toc168515995)

[BẢNG PHÂN CÔNG NHIỆM VỤ 4](#_Toc168515996)

[MỞ ĐẦU 6](#_Toc168515997)

[CHƯƠNG 1 GIỚI THIỆU VỀ DEEPFAKE 8](#_Toc168515998)

[1.1 Định nghĩa Deepfake và cách hoạt động: 8](#_Toc168515999)

[1.1.1 Định nghĩa: 8](#_Toc168516000)

[1.1.2 Cách hoạt động: 8](#_Toc168516001)

[1.2 Lịch sử và phát triển của công nghệ Deepfake: 8](#_Toc168516002)

[1.3 Ứng Dụng Phổ Biến của Deepfake: 9](#_Toc168516003)

[CHƯƠNG 2 LỢI ÍCH VÀ TÁC HẠI DEEPFAKE 10](#_Toc168516004)

[2.1 Lợi ích của Deepfake: 10](#_Toc168516005)

[2.2 Nguy Hại của Deepfake: 10](#_Toc168516006)

[2.2.1 Lừa Đảo và Lan Truyền Thông Tin Sai Lệch: 10](#_Toc168516007)

[2.2.2 Sản Phẩm Đồi Trụy: 10](#_Toc168516008)

[2.2.3 Tấn Công Phi Kỹ Thuật và Lừa Đảo: 10](#_Toc168516009)

[CHƯƠNG 3 CÁC VẤN ĐỀ ĐẠO ĐỨC LIÊN QUAN ĐẾN DEEPFAKE VÀ ẢNH HƯỞNG TRONG THÔNG TIN VÀ TRUYỀN THÔNG 11](#_Toc168516010)

[3.1 Thông tin sai lệch và thao túng: 11](#_Toc168516011)

[3.2 Vi phạm quyền riêng tư: 11](#_Toc168516012)

[3.3 Thiếu tính xác thực: 12](#_Toc168516013)

[CHƯƠNG 4 ẢNH HƯỞNG CỦA DEEPFAKE ĐỐI VỚI XÃ HỘI 13](#_Toc168516014)

[4.1 Tác động của Deepfake đối với quan hệ giữa cá nhân và xã hội. Hậu quả 13](#_Toc168516015)

[4.1.1 Tác động của Deepfake đối với quan hệ giữa cá nhân và xã hội 13](#_Toc168516016)

[4.1.2 Hậu quả 14](#_Toc168516017)

[4.2 Nguy cơ về sự tin tưởng và ảnh hưởng của thông tin sai lệch hoặc giả mạo 15](#_Toc168516018)

[4.2.1 Thông tin sai lệch là gì? 15](#_Toc168516019)

[4.2.2 Nguy cơ về sự tin tưởng 15](#_Toc168516020)

[4.2.3 Ảnh hưởng của thông tin sai lệch hoặc giả mạo 15](#_Toc168516021)

[4.3 Tác động của Deepfake đối với các lĩnh vực như chính trị, ngành công nghiệp giải trí, quảng cáo và hậu quả 17](#_Toc168516022)

[4.3.1 Tác động đến chính trị 17](#_Toc168516023)

[CHƯƠNG 5 VÍ DỤ THỰC TẾ VỀ VIỆC SỬ DỤNG DEEPFAKE ĐỂ GIẢ MẠO 22](#_Toc168516024)

[5.1 Quy trình để tạo ra Deepfake giả mạo khuôn mặt người khác 22](#_Toc168516025)

[5.1.1 Thu thập dữ liệu 22](#_Toc168516026)

[5.1.2 Chuẩn bị dữ liệu 22](#_Toc168516027)

[5.1.3 Huấn luyện mô hình 23](#_Toc168516028)

[5.1.4 Tạo deepfake 24](#_Toc168516029)

[5.1.5 Kiểm tra và hoàn thiện 25](#_Toc168516030)

[5.2 Công cụ và phần mềm phát triển 26](#_Toc168516031)

[CHƯƠNG 6 CÁC BIỆN PHÁP PHÒNG NGỪA VÀ GIẢI QUYẾT 27](#_Toc168516032)

[CHƯƠNG 7 PHÂN TÍCH 29](#_Toc168516040)

[7.1 Phân tích các nghiên cứu hiện có và thảo luận trong cộng đồng học thuật và chính trị về ảnh hưởng của Deepfake: 29](#_Toc168516041)

[7.2 So sánh với các công nghệ và hiện tượng tương tự khác như tin tức giả mạo, sự thay đổi hình ảnh,… 33](#_Toc168516042)

[CHƯƠNG 8 KẾT LUẬN VÀ ĐỀ XUẤT 35](#_Toc168516043)

[8.1 Tóm tắt nhấn mạnh tầm quan trọng của vấn đề 35](#_Toc168516044)

[8.2 Đề xuất các biện pháp cụ thể để ứng phó với ảnh hưởng của Deepfake đối với xã hội và đạo đức 35](#_Toc168516045)

[CHƯƠNG 9 CÁC VÍ DỤ VỀ LỢI ÍCH CỦA DEEPFAKE TỪ TRƯỚC TỚI NAY. 39](#_Toc168516046)

[9.1 Bùng nổ nhu cầu dùng deepfake nói chuyện với người đã khuất 39](#_Toc168516047)

[9.2 Deepfake triển khai cho digital marketing 40](#_Toc168516048)

[9.3 Deepfake ứng dụng trong giáo dục: 40](#_Toc168516049)

[9.4 Deepfake dùng trong làm film 41](#_Toc168516050)

[CHƯƠNG 10 CÁC VÍ DỤ VỀ HẬU QUẢ CỦA DEEPFAKE TỪ TRƯỚC TỚI NAY 42](#_Toc168516051)

[10.1 Video tổng thống obama được tạo bởi deepfake 42](#_Toc168516052)

[10.2 Các sản phẩm đồi trụ có gương mặt người nổi tiếng 42](#_Toc168516053)

[10.3 Cuộc gọi video Deepfake để lừa đảo 44](#_Toc168516054)

[KẾT LUẬN 45](#_Toc168516055)

[Tài liệu tham khảo 47](#_Toc168516056)

# 

# DANH MỤC CÁC HÌNH VẼ

[Hình 5‑1 21](#_Toc168515782)

[Hình 5‑2 22](#_Toc168515783)

[Hình 5‑3 23](#_Toc168515784)

[Hình 5‑4 25](#_Toc168515785)

[Hình 9‑1 39](#_Toc168515786)

[Hình 9‑2 40](#_Toc168515787)

[Hình 10‑1 41](#_Toc168515788)

[Hình 10‑2 42](#_Toc168515789)

[Hình 10‑3 43](#_Toc168515790)

# DANH MỤC CÁC BẢNG BIỂU

[BẢNG PHÂN CÔNG NHIỆM VỤ 5](#_Toc167910451)

# BẢNG PHÂN CÔNG NHIỆM VỤ

| STT | Thành viên | Ảnh | Công việc | Tình trạng |
| --- | --- | --- | --- | --- |
| 1 | Bùi Đức Thắng  (Nhóm trưởng)  2121050756 |  | - Các ví dụ về lợi ích của deepfake từ trước tới nay.  - Các ví dụ về hậu quả của deepfake từ trước tới nay.  - Ví dụ thực tế về việc sử dụng deepfake để giả mạo.  - Làm powpoint | 100%  100%  100%  100% |
| 2 | Nguyễn Đức Thành  2121051158 |  | - Ví dụ thực tế về việc sử dụng deepfake để giả mạo.  - Các biện pháp phòng ngừa và giả quyết.  - Làm word. | 100%  100%  100% |
| 3 | Đinh Việt Hưng  2121050066 |  | - Các vấn đề đạo đức liên quan đến deepfake và ảnh hưởng trong thông tin và truyền thông.  - Ảnh hưởng của deepfake đối với xã hội. | 100%  100% |
| 4 | Nguyễn Quang Đạo  2121050451 |  | - Giới thiệu về Deepfake.  - Phân tích lợi ích và tác hại Deepfake. | 100%  100% |
| 5 | Đinh Quang Trọng  2121050483 |  | - Phân tích.  - Kết luận và đề xuất | 100%  100% |

# MỞ ĐẦU

Công nghệ giả mạo sâu (deepfake) đã trở thành một hiện tượng nổi bật trong kỷ nguyên số hóa, mang lại những tiềm năng vượt bậc nhưng cũng đặt ra nhiều thách thức phức tạp đối với xã hội và đạo đức. Deepfake sử dụng các kỹ thuật học sâu để tạo ra những hình ảnh, video, và âm thanh giả mạo có độ chân thực cao, mở ra những cơ hội mới trong các lĩnh vực như giải trí, quảng cáo và giáo dục. Tuy nhiên, song song với những lợi ích này, deepfake cũng đem lại những nguy cơ nghiêm trọng.

Một trong những tác động tiêu cực đáng lo ngại nhất của deepfake là khả năng lan truyền thông tin sai lệch và lừa đảo. Các video giả mạo của các nhà lãnh đạo chính trị, người nổi tiếng, hoặc bất kỳ cá nhân nào có thể được sử dụng để phát tán thông tin sai lệch, gây rối loạn xã hội, hoặc thậm chí là gây ra các cuộc xung đột. Ngoài ra, deepfake còn có thể được sử dụng để tống tiền, bôi nhọ danh tiếng hoặc xâm phạm quyền riêng tư của cá nhân, gây ra những tổn thất nặng nề về mặt tinh thần và tài chính.

Sự khó khăn trong việc phân biệt giữa thật và giả mà deepfake gây ra còn dẫn đến một vấn đề nghiêm trọng hơn: sự suy giảm niềm tin vào truyền thông và các nguồn thông tin trực tuyến. Khi công chúng không còn tin tưởng vào những gì họ thấy và nghe, điều này có thể gây ra một trạng thái hoang mang và mất niềm tin rộng rãi, làm ảnh hưởng đến sự ổn định của xã hội và sự tin cậy trong các mối quan hệ cá nhân.

Từ góc độ đạo đức, deepfake đặt ra nhiều câu hỏi phức tạp về việc sử dụng và lạm dụng công nghệ. Việc sử dụng hình ảnh và giọng nói của người khác mà không có sự đồng ý rõ ràng không chỉ xâm phạm quyền riêng tư mà còn vi phạm các nguyên tắc đạo đức cơ bản về sự tôn trọng và sự đồng ý cá nhân. Các quy định pháp lý hiện tại ở nhiều quốc gia vẫn chưa theo kịp sự phát triển nhanh chóng của công nghệ này, đòi hỏi các nhà làm luật phải nhanh chóng xây dựng và cập nhật các quy định để bảo vệ quyền lợi của công dân và duy trì trật tự xã hội.

Để đối phó với những thách thức mà deepfake mang lại, cần có sự hợp tác chặt chẽ giữa các nhà khoa học, các công ty công nghệ, và các cơ quan quản lý. Nghiên cứu và phát triển các công cụ để phát hiện và ngăn chặn deepfake là cần thiết hơn bao giờ hết. Đồng thời, việc giáo dục cộng đồng về cách nhận diện thông tin giả mạo và tăng cường nhận thức về đạo đức trong việc sử dụng công nghệ cũng là những bước quan trọng nhằm giảm thiểu tác động tiêu cực của deepfake.

Công nghệ deepfake, với những tiềm năng và thách thức của nó, đòi hỏi một cách tiếp cận cân bằng và có trách nhiệm. Chỉ khi chúng ta hiểu rõ và kiểm soát được công nghệ này, chúng ta mới có thể khai thác tối đa những lợi ích mà nó mang lại đồng thời bảo vệ xã hội khỏi những hệ lụy tiêu cực.

# GIỚI THIỆU VỀ DEEPFAKE

## Định nghĩa Deepfake và cách hoạt động:

### Định nghĩa:

+Deepfake là kết hợp giữa “deep learning” (học sâu) và “fake” (giả mạo). Deepfake sử dụng trí tuệ nhân tạo và máy học để tạo nội dung âm thanh và hình ảnh giả mạo có thể dễ dàng đánh lừa.

### Cách hoạt động:

+ Deepfake quét dữ liệu video và ảnh chân dung của một người, sau đó sử dụng trí tuệ nhân tạo để thay thế các chi tiết như mắt, miệng, mũi, tạo ra gương mặt, giọng nói giống như thật. Công nghệ này có thể tạo ra những video lừa đảo chân thực khó nhận ra.

## Lịch sử và phát triển của công nghệ Deepfake:

- Bước đầu tiên (trước năm 2017):

+ Trước khi thuật ngữ "deepfake" trở nên phổ biến, đã có nhiều dự án và nghiên cứu về việc tạo ra hình ảnh giả mạo bằng các phương pháp học máy và học sâu.

+ Các phần mềm như FaceSwap đã xuất hiện, cho phép người dùng hoán đổi khuôn mặt giữa các video một cách tương đối đơn giản.

- Sự xuất hiện của Deepfake (2017):

+ Deepfake trở thành một từ khóa nổi bật sau khi một số video giả mạo nổi tiếng bắt đầu lan truyền trên Internet.

+ Công nghệ deepfake ban đầu được sử dụng để tạo ra các video khiêu dâm giả mạo với khuôn mặt của các nữ diễn viên nổi tiếng.

- Sự tiến bộ (2018-2020):

+ Các nghiên cứu và phát triển trong lĩnh vực deep learning và computer vision đã làm cho công nghệ deepfake trở nên mạnh mẽ hơn và dễ dàng sử dụng hơn.

+ Các công cụ và phần mềm deepfake dễ sử dụng đã được phát triển, làm cho việc tạo ra video deepfake trở nên đơn giản hơn đối với người dùng không chuyên.

- Cảnh báo và phản ứng (từ 2020 trở đi):

+ Sự phổ biến của deepfake đã gây ra nhiều lo ngại về vấn đề an ninh, quyền riêng tư và sự tin tưởng trong các lĩnh vực như chính trị, giải trí và xã hội.

+ Các nghiên cứu và các tổ chức bảo vệ quyền riêng tư đã bắt đầu nghiên cứu cách phát hiện và ngăn chặn sự lạm dụng của công nghệ deepfake.

- Phát triển trong việc xử lý (2020 - nay):

+ Các phương pháp và công nghệ mới tiếp tục được phát triển để giúp phát hiện và xử lý các video và hình ảnh deepfake.

+ Các tổ chức và chính phủ cũng đang thảo luận về cách điều chỉnh và quản lý việc sử dụng deepfake trong các hoạt động trái pháp luật và nguy hiểm cho cộng đồng.

## Ứng Dụng Phổ Biến của Deepfake:

**- Giải trí**: Tạo video giải trí bằng việc đổi mặt các diễn viên, biến mình thành người nổi tiếng.

**- Chính trị**: Sử dụng để tạo video giả mạo người nổi tiếng với phát ngôn bất bình.

**- Xã hội**: Lợi dụng để lừa đảo, chiếm đoạt tài sản thông qua video và âm thanh giả mạo.

# 

# LỢI ÍCH VÀ TÁC HẠI DEEPFAKE

## Lợi ích của Deepfake:

- Deepfake cho phép thương hiệu tiết kiệm chi phí sản xuất video quảng cáo bằng cách sử dụng công nghệ để tạo nội dung chất lượng cao mà không cần đến diễn viên người đóng.

- Doanh nghiệp có thể sử dụng hình ảnh cũ hoặc mua giấy phép sử dụng hình ảnh diễn viên để tạo ra sản phẩm quảng cáo mới.

## Nguy Hại của Deepfake:

### Lừa Đảo và Lan Truyền Thông Tin Sai Lệch:

- Deepfake có thể được sử dụng để tạo nội dung lừa đảo, lan truyền thông tin sai lệch, và ảnh hưởng nghiêm trọng đến tâm trí của người xem.

- Có đến 96% video sử dụng Deepfake chứa nội dung đồi truỵ và có thể gây ra hậu quả không lường trước.

### Sản Phẩm Đồi Trụy:

- Deepfake có thể tạo ra video khiêu dâm giả mạo, tin tức giả mạo, và lừa đảo, gây tổn thất nghiêm trọng cho nạn nhân.

- Nguy cơ tống tiền và gia tăng mối đe dọa từ các vấn đề như revenge porn được tăng lên.

### Tấn Công Phi Kỹ Thuật và Lừa Đảo:

- Deepfake có thể được sử dụng trong tấn công phi kỹ thuật để dụ người dùng tiết lộ thông tin bí mật hoặc gây ra các trò lừa đảo khác.

- Khả năng tạo ra video và hình ảnh giả mạo gần như không thể phân biệt, dẫn đến việc phủ nhận các bằng chứng có thật.

- Trong tổng thể, Deepfake mang lại cơ hội sáng tạo và tiện lợi trong nhiều lĩnh vực, nhưng cũng đồng thời đặt ra các vấn đề nghiêm trọng về an ninh mạng, quyền riêng tư, và đạo đức thông tin. Để đối phó, cần có biện pháp phòng ngừa và kiểm soát chặt chẽ đối với sự lạm dụng của công nghệ này.

# CÁC VẤN ĐỀ ĐẠO ĐỨC LIÊN QUAN ĐẾN DEEPFAKE VÀ ẢNH HƯỞNG TRONG THÔNG TIN VÀ TRUYỀN THÔNG

## Thông tin sai lệch và thao túng:

**- Lan truyền thông tin sai lệch và tin giả**: Deepfake có thể được sử dụng để tạo ra những nội dung giả mạo một cách thuyết phục, khiến người xem khó phân biệt được với thông tin thực tế. Điều này có thể dẫn đến việc lan truyền thông tin sai lệch và tin giả, gây nhiễu loạn thông tin và thao túng dư luận.

**- Làm suy giảm niềm tin vào truyền thông**: Khi người xem không thể chắc chắn về tính xác thực của những gì họ nhìn thấy và nghe thấy, họ có thể bắt đầu nghi ngờ tất cả các nguồn tin tức, dẫn đến sự sụt giảm niềm tin vào truyền thông nói chung.

**- Gây tổn hại đến danh tiếng**: Deepfake có thể được sử dụng để tạo ra những video hoặc bản ghi âm giả mạo nhằm làm hỏng danh tiếng của cá nhân hoặc tổ chức. Điều này có thể gây ra hậu quả nghiêm trọng về mặt cá nhân và chuyên nghiệp.

**- Gây khó khăn cho việc thực thi pháp luật**: Deepfake có thể được sử dụng để tạo bằng chứng giả mạo hoặc che giấu các hành vi sai trái. Điều này có thể làm cản trở việc thực thi pháp luật và truy tố tội phạm.

## Vi phạm quyền riêng tư:

- Deepfake cũng có thể được sử dụng để vi phạm quyền riêng tư của mọi người. Ví dụ, một deepfake có thể được tạo ra để hiển thị một ai đó đang làm điều gì đó mà họ không bao giờ làm, chẳng hạn như tham gia vào một hành vi khiêu dâm hoặc phạm tội. Điều này có thể gây ra tổn hại đáng kể đến danh tiếng và sự an toàn của họ.

- Việc tạo ra deepfake thường đòi hỏi việc sử dụng hình ảnh và video của mọi người mà không có sự đồng ý của họ. Điều này vi phạm quyền riêng tư của họ và có thể khiến họ cảm thấy bị bóc lột hoặc bị lạm dụng.

## Thiếu tính xác thực:

- Deepfake có thể rất thuyết phục và khó phân biệt với video hoặc bản ghi âm thực tế. Điều này có thể khiến mọi người khó tin tưởng vào những gì họ nhìn thấy và nghe thấy, dẫn đến sự gia tăng hoài nghi và ngờ vực.

- Deepfake có thể được sử dụng cho nhiều mục đích bất hợp pháp, chẳng hạn như tống tiền, gian lận hoặc lừa đảo. Điều này có thể gây ra thiệt hại về tài chính và tinh thần đáng kể cho các nạn nhân.

# 

# ẢNH HƯỞNG CỦA DEEPFAKE ĐỐI VỚI XÃ HỘI

## Tác động của Deepfake đối với quan hệ giữa cá nhân và xã hội. Hậu quả

### Tác động của Deepfake đối với quan hệ giữa cá nhân và xã hội

- Deepfake có thể trở thành mối đe dọa đáng lo ngại đối với xã hội và cá nhân. Công nghệ này có thể được sử dụng để làm hoen ố danh tiếng của người nổi tiếng hoặc tạo video giả mạo. Điều này có thể dẫn đến nhiều hậu quả tiêu cực cho đối tượng, cả về tinh thần lẫn tài chính.

- Ngoài ra, deepfake có thể được sử dụng để phát tán tin giả, làm xói mòn lòng tin của công chúng và gây ra những hậu quả đáng tiếc cho toàn xã hội. Những video giả mạo như thế này có thể gây nhầm lẫn, khiến người xem khó phân biệt được đâu là video thật và video giả.

- Công nghệ deepfake không chỉ gây hại cho đối tượng mà còn là mối đe dọa đáng kể đối với an ninh quốc gia. Nó có thể được sử dụng để tạo ra các video sai sự thật về các nguyên thủ quốc gia và các quan chức cấp cao khác, gây khó khăn, đe dọa quan hệ quốc tế.

- Nghiên cứu gần đây cũng cho thấy deepfake ngày càng phổ biến trên mạng xã hội. Bất cứ ai đều có thể dễ dàng tạo ra những video giả mạo chỉ bằng vài cú nhấp chuột và sử dụng những công cụ phổ biến như Photoshop, FaceSwap hay DeepFaceLab. Hơn nữa, các video giả mạo đang trở nên ngày càng chân thực hơn và khó phân biệt với sự thật. Điều này làm cho nhiều người cảm thấy lo ngại về tương lai của mình trên mạng xã hội. [1]

- Deepfake có thể tác động mạnh mẽ đến quan hệ giữa cá nhân và xã hội từ nhiều khía cạnh khác nhau:

+ Các mối đe dọa đối với quyền riêng tư: Deepfake có khả năng làm suy yếu quyền riêng tư của một cá nhân vì họ có thể tham gia vào các video không phù hợp hoặc thậm chí tham gia vào các hoạt động mà họ không tham gia.

+ Uy tín và danh tiếng bị đe dọa: Nếu một người bị sản xuất video hoặc hình ảnh deepfake xấu, uy tín của họ trong cộng đồng xã hội và các môi trường chuyên môn có thể bị tổn hại nghiêm trọng.

+ Tác động đến tin tức và thông tin: Deepfake có thể làm xói mòn niềm tin của mọi người vào tin tức và thông tin trực tuyến bằng cách làm xáo trộn sự thật với sự giả dối.

+ Tăng nguy cơ lạm dụng và vi phạm: Video deepfake có thể được sử dụng để đánh lừa hoặc lạm dụng người khác, chẳng hạn như sử dụng video giả mạo để gây tranh cãi hoặc làm tổn hại đến danh tiếng của một cá nhân hoặc vị trí tổ chức.

+ Gây hoang mang và sợ hãi trong cộng đồng: Sự phổ biến của deepfake có thể làm tăng nỗi sợ hãi và lo lắng của cộng đồng về việc mất thông tin cá nhân cũng như độ tin cậy của các phương tiện truyền thông.

### Hậu quả

- Hậu quả của sự lan rộng của deepfake có thể rất nghiêm trọng và đa dạng, bao gồm:

+ Thiệt hại cho cá nhân, tổ chức: Khi sử dụng video deepfake để mạo danh hoặc lừa đảo, danh tiếng, uy tín và tài sản của cá nhân, tổ chức có thể bị tổn hại.

+ Mất niềm tin vào thông tin: Sự lan truyền của deepfake có thể làm xói mòn niềm tin vào thông tin trực tuyến, gây hoang mang và lo ngại trong cộng đồng.

+ Gây tranh cãi, xung đột: Các video deepfake có thể được sử dụng để tạo ra tin giả hoặc gây tranh cãi, xung đột giữa các cá nhân, tổ chức.

+ Mất niềm tin vào các phương tiện truyền thông chính thống và chính phủ: Nếu deepfake lan truyền rộng rãi, nó có thể dẫn đến mất lòng tin vào các phương tiện truyền thông chính thống và các tổ chức chính phủ.

+ Tăng nguy cơ xâm nhập và lừa đảo: Deepfake có thể được sử dụng để thực hiện hành vi lừa đảo hoặc đột nhập vào hệ thống máy tính hoặc tài khoản trực tuyến của người dùng.

+ Làm trầm trọng thêm sự phân biệt đối xử và căng thẳng xã hội: Deepfake có thể làm trầm trọng thêm sự phân biệt đối xử và căng thẳng xã hội bằng cách truyền bá những thông điệp thù địch hoặc phân biệt chủng tộc, giới tính, văn hóa.

## Nguy cơ về sự tin tưởng và ảnh hưởng của thông tin sai lệch hoặc giả mạo

### Thông tin sai lệch là gì?

- Thông tin sai lệch là tin tức, câu chuyện hoặc trò lừa bịp được tạo ra nhằm mục đích cố tình đánh lừa hoặc đánh lừa người đọc. Thông thường mục đích của những câu chuyện này là gây ảnh hưởng đến quan điểm của mọi người, ảnh hưởng đến chương trình nghị sự chính trị hoặc gây nhầm lẫn và thường là một hoạt động kinh doanh sinh lợi cho các nhà xuất bản trực tuyến. Thông tin sai lệch có thể đánh lừa mọi người bằng cách trông giống như một trang web đáng tin cậy hoặc sử dụng tên và URL tương tự như tên và URL của một tổ chức tin tức có uy tín. Chuyên gia truyền thông Martina Chapman cho biết ba yếu tố dẫn đến tin giả: "nghi ngờ, thông tin sai lệch và thao túng". [2]

### Nguy cơ về sự tin tưởng

- Tin tưởng vào thông tin và tin tức: Deepfake làm suy yếu độ tin cậy của thông tin và tin tức trực tuyến. Khi một video hoặc hình ảnh được tạo ra dễ dàng và khó phân biệt với hàng thật, người tiêu dùng sẽ khó phân biệt đâu là thật đâu là giả.

- Niềm tin vào người khác: Deepfake có thể làm suy yếu niềm tin vào người khác khi các cá nhân hoặc tổ chức bị mạo danh trong video hoặc hình ảnh giả mạo.

- Niềm tin vào truyền thông và chính phủ: Nếu deepfake trở nên phổ biến, nó có thể dẫn đến mất niềm tin vào các phương tiện truyền thông chính thống và các tổ chức chính phủ, vì mọi người không còn tin tưởng vào tính chính xác và độ tin cậy của thông tin họ cung cấp.

- Niềm tin vào công nghệ: Sự phổ biến của deepfake có thể dẫn đến mất lòng tin vào công nghệ, vì người dùng lo ngại rằng việc sử dụng công nghệ có thể gây ra hậu quả tiêu cực và gian lận.

- Tin tưởng vào bản thân: Deepfake có thể gây ra sự nghi ngờ và lo lắng về bản thân khi mỗi người đều có thể trở thành nạn nhân của việc mạo danh hoặc bị lừa đảo thông qua video hoặc hình ảnh giả mạo.

### Ảnh hưởng của thông tin sai lệch hoặc giả mạo

- Ảnh hưởng đến chính trị:

+ Thay đổi quyết định của cử tri: Thông tin sai lệch có thể thay đổi quyết định của cử tri bằng cách thay đổi hoặc bóp méo thông tin về ứng cử viên, chính sách và các vấn đề chính trị.

+ Gây bất ổn, suy yếu an ninh chính trị: Tin giả có thể được sử dụng để gây bất ổn, suy yếu an ninh chính trị và tạo ra hỗn loạn, xung đột trong xã hội.

+ Gây chia rẽ và căng thẳng: Có thể làm gia tăng sự chia rẽ và căng thẳng trong xã hội bằng cách truyền bá những thông tin sai lệch, thù hận hoặc bè phái.

+ Thay đổi hệ thống chính trị: Việc truyền bá thông tin sai lệch có thể làm suy yếu hệ thống chính trị bằng cách gây mất lòng tin vào các thể chế chính phủ và các đảng phái chính trị.

+ Tạo ra tranh cãi và hỗn loạn: Thông tin sai lệch có thể tạo ra tranh cãi và hỗn loạn trong xã hội bằng cách truyền bá những thông điệp có hại và gây nhầm lẫn.

+ Gây hại cho đất nước và thế giới: Thông tin sai lệch có thể gây hại cho đất nước và thế giới bằng cách làm suy yếu quan hệ đối ngoại và tạo ra sự bất ổn trên thế giới. [2]

- Ảnh hưởng đến kinh tế

+ Tác động đến quyết định đầu tư và kinh doanh: Thông tin sai lệch có thể hủy hoại niềm tin của nhà đầu tư và doanh nghiệp, tạo ra sự không chắc chắn và lo lắng về tương lai của công ty. Điều này có thể dẫn đến giảm đầu tư và tiêu dùng, có thể ảnh hưởng tiêu cực đến tăng trưởng kinh tế.

+ Biến động trên thị trường tài chính: Tin giả có thể gây ra biến động trên thị trường tài chính bằng cách thay đổi quan điểm và kỳ vọng của nhà đầu tư về điều kiện kinh tế và kinh doanh.

+ Tạo ra sự không chắc chắn và rủi ro: Thông tin sai lệch tạo ra các hoạt động kinh doanh không chắc chắn và rủi ro, khiến các công ty đưa ra những quyết định khó khăn và đối mặt với sự không chắc chắn về chiến lược kinh doanh của họ.

+ Gây bất ổn và sợ hãi: Thông tin sai lệch có thể gây ra bất ổn và lo sợ cho nền kinh tế, khiến nhà đầu tư và doanh nghiệp thận trọng hơn, trì hoãn các quyết định đầu tư và tăng trưởng kinh doanh.

+ Tác hại đối với hệ thống tài chính: Thông tin sai lệch có thể gây tổn hại cho hệ thống tài chính bằng cách làm suy yếu niềm tin của người tiêu dùng và nhà đầu tư, tạo ra sự mất cân bằng và bất ổn trên thị trường. [2].

- Ảnh hưởng đến xã hội

+ Gây chia rẽ và căng thẳng: Thông tin sai lệch có thể làm tăng sự chia rẽ và căng thẳng trong xã hội bằng cách truyền bá những thông điệp thù hận hoặc phân biệt đối xử dựa trên chủng tộc, tôn giáo, giới tính hoặc văn hóa.

+ Gây mất lòng tin và nhầm lẫn: Thông tin sai lệch làm mất niềm tin vào giới truyền thông và chính phủ, đồng thời tạo ra sự nhầm lẫn và sợ hãi trong công chúng.

+ Phân biệt đối xử và kích động: Thông tin sai lệch có thể được sử dụng để truyền bá các thông điệp mang tính kích động hoặc thúc đẩy sự phân biệt đối xử dựa trên các đặc điểm như tôn giáo, chính trị hoặc quốc tịch.

+ Thiệt hại về niềm tin vào hệ thống: Thông tin sai lệch có thể làm suy yếu niềm tin về chính trị, pháp lý và xã hội, làm suy yếu an ninh và làm tăng sự sợ hãi và mất lòng tin.

+ Gây tranh cãi và xung đột: Thông tin sai lệch có thể gây ra xung đột và xung đột trong xã hội bằng cách truyền bá những thông điệp có hại và gây nhầm lẫn cũng như làm gia tăng sự phân biệt đối xử và bất đồng chính kiến.

+ Làm tổn hại đến sự hợp tác và hòa hợp: Thông tin sai lệch có thể làm suy yếu sự hợp tác và hòa hợp trong xã hội bằng cách làm mất niềm tin giữa các đồng nghiệp và dẫn đến xung đột và chia rẽ. [2]

## Tác động của Deepfake đối với các lĩnh vực như chính trị, ngành công nghiệp giải trí, quảng cáo và hậu quả

### Tác động đến chính trị

- Làm suy yếu niềm tin vào chính trị: Deepfakes có thể làm suy yếu niềm tin của công chúng đối với các nhà lãnh đạo chính trị và tổ chức chính phủ, vì hình ảnh hoặc video giả mạo có thể gây hiểu lầm về họ.

- Gây tranh cãi và bất ổn: Deepfakes có thể gây tranh cãi và bất ổn chính trị bằng cách truyền bá tin tức giả mạo hoặc phản đối, tạo ra căng thẳng và bất ổn.

- Hỗ trợ các chiến dịch quảng cáo và pháp lý: Deepfake có thể được sử dụng để quảng bá các chiến dịch quảng cáo và pháp lý bằng cách tạo ra hình ảnh hoặc video giả mạo để gây ảnh hưởng đến dư luận.

- Làm suy yếu tính minh bạch và liêm chính: Deepfakes có thể làm suy yếu tính minh bạch và liêm chính chính trị bằng cách xuất bản thông tin sai lệch hoặc mâu thuẫn, gây nhầm lẫn và nghi ngờ về sự thật.

- Đe dọa an ninh quốc gia: Deepfake có thể gây ra mối đe dọa nghiêm trọng cho an ninh quốc gia bằng cách tạo ra các hình ảnh hoặc video giả gây gián đoạn hoặc phá hủy.

- Tạo ra mối lo ngại và lo lắng: Deepfake có thể gây lo ngại và lo ngại của công chúng về việc công nghệ được sử dụng cho mục đích lừa đảo hoặc gây ra tác động tiêu cực về mặt chính trị.

Tác động đến ngành công nghiệp giải trí

- Tin nhắn giả mạo trong nghệ thuật và giải trí: Deepfakes có thể được sử dụng để tạo video hoặc ảnh giả mạo của những người nổi tiếng, gây nhầm lẫn và tranh luận về tính xác thực của nội dung.

- Quyền sở hữu trí tuệ: Deepfake có thể thách thức quyền sở hữu trí tuệ khi sử dụng hình ảnh và video của người khác mà không có sự cho phép của họ.

- Tác động đến sự thực tế và đạo đức trong giải trí: Sự phổ biến của deepfake có thể làm mất đi sự thực tế và đạo đức trong ngành công nghiệp giải trí bằng cách tạo ra các hình ảnh và video giả mạo không phản ánh đúng hiện thực.

- Phát triển nâng cao : Deepfake cũng có thể nâng cao ngành công nghiệp giải trí bằng cách tạo ra các hiệu ứng và hiệu ứng đặc biệt, mang lại cơ hội cho sự sáng tạo và đổi mới.

- Làm tăng mối lo ngại về quyền riêng tư: Deepfake có thể gây lo ngại về quyền riêng tư bằng cách sử dụng hình ảnh và video của người khác mà không có sự đồng ý của họ, gây tổn hại nghiêm trọng cho cá nhân và gia đình họ.

Tác động đến ngành quảng cáo

- Tăng khả năng sáng tạo nội dung hấp dẫn: Deepfake có thể được sử dụng để tạo video, video quảng cáo tạo hiệu ứng đặc biệt và thu hút nhiều người xem hơn. Việc sử dụng các lớp sâu có thể mở ra nhiều khả năng sáng tạo khi tạo sản phẩm quảng cáo.

- Mất niềm tin và sự tín nhiệm: Việc giả mạo sâu cũng có thể làm suy yếu niềm tin của người tiêu dùng vào tiếp thị nội dung. Nếu người tiêu dùng cảm thấy quảng cáo có thể bị thao túng hoặc lừa dối, họ có thể không tin tưởng vào thông điệp được truyền tải.

- Tăng nguy cơ khai thác dữ liệu: Deepfake có thể được sử dụng để tạo video quảng cáo giả mạo, gây nhầm lẫn và dễ đánh lừa người xem. Điều này có thể làm tăng nguy cơ lạm dụng dữ liệu và gây hậu quả tiêu cực cho các thương hiệu và công ty.

- Phản đối quyền sở hữu trí tuệ: Khi bạn sử dụng hình ảnh và video của người khác mà không có sự đồng ý hoặc cho phép của họ, điều đó có thể gây ra các vấn đề về quyền sở hữu trí tuệ.

- Nhu cầu về công nghệ nhận dạng giả mạo sâu ngày càng tăng: Trước nguy cơ lạm dụng phổ biến trong quảng cáo, ngành quảng cáo có thể tăng nhu cầu về công nghệ nhận dạng giả mạo sâu để duy trì niềm tin và tính toàn vẹn.

Hậu quả

- Hậu quả đối với chính trị

+ Mất niềm tin và uy tín: Deepfake có thể gây mất niềm tin và uy tín của các nhà lãnh đạo và tổ chức chính phủ khi tin tức giả lan truyền.

+ Tranh cãi và căng thẳng gia tăng: Deepfake có thể làm tăng tranh cãi và căng thẳng chính trị bằng cách truyền bá các thông điệp hoặc phản đối đe dọa.

+ Khủng hoảng quyền lực và ổn định: Sự trỗi dậy của deepfake có thể tạo ra sự bất ổn và hoang mang trong dân chúng, gây ra khủng hoảng quyền lực và bất ổn chính trị.

- Hậu quả đối với ngành giải trí

+ Mất niềm tin vào tin tức và nghệ sĩ: Deepfake có thể gây mất niềm tin vào tin tức và nghệ sĩ khi tin giả lan truyền, khiến khán giả không tin vào tính xác thực của giải trí.

+ Thiệt hại đến uy tín của ngành: Lạm dụng deepfake có thể làm tổn hại đến uy tín của ngành giải trí và dẫn đến mất niềm tin của khán giả và nhà đầu tư.

- Hậu quả đối với ngành quảng cáo

+ Mất niềm tin của khách hàng: Deepfake có thể khiến khách hàng mất niềm tin vào các thông điệp tiếp thị và thương hiệu khi thông tin sai lệch được sử dụng để quảng cáo sản phẩm hoặc dịch vụ.

+ Thiệt hại về thương hiệu: Việc lạm dụng deepfake quá mức trong quảng cáo có thể làm hỏng thương hiệu do ảnh hưởng đến độ tin cậy và tính toàn vẹn của thông điệp quảng cáo.

Tác động của deepfake với quyền riêng tư và an ninh cá nhân và hậu quả

Tác động của deepfake với quyền riêng tư và an ninh cá nhân

- Vi phạm quyền riêng tư: Deepfake có thể được sử dụng để tạo video và hình ảnh giả mạo của ai đó mà không được phép. Điều này có thể dẫn đến mất quyền riêng tư, mất niềm tin và tổn hại đến danh tiếng của người bị giả mạo.

- Gian lận và gián đoạn: Deepfake có thể được sử dụng để tạo thông tin sai lệch, chẳng hạn như video hoặc tin nhắn giả mạo, nhằm tạo ra một sự gian lận hoặc gián đoạn. Điều này có thể gây ra hậu quả nghiêm trọng, từ gian lận tài chính đến bóp méo thông tin trong các cuộc tranh luận chính trị.

- Nguy cơ gia tăng đối với người nổi tiếng: Những người nổi tiếng, chẳng hạn như chính trị gia hoặc người nổi tiếng trong làng giải trí, có thể trở nên dễ bị tổn thương hơn. Việc tạo ra những hình ảnh giả mạo có thể gây ra hậu quả nghiêm trọng, từ hủy hoại sự nghiệp đến tạo ra những tình huống xã hội tiêu cực.

- Gây hỗn loạn và nhầm lẫn: Deepfake có thể được sử dụng để tạo ra hình ảnh hoặc thông tin giả mạo về các chủ đề quan trọng hoặc những người quan trọng. Sự lan truyền của thông tin sai lệch này có thể gây ra sự hỗn loạn và bất ổn trong xã hội.

- Đe dọa quyền riêng tư: Deepfake có thể được sử dụng để đánh lừa hoặc bôi nhọ mọi người bằng cách tạo các video giả mạo cho thấy họ tham gia vào các hoạt động hoặc hoạt động không phù hợp.

Hậu quả

- Thiệt hại về danh tiếng và niềm tin: Deepfake có thể gây tổn hại đến danh tiếng và niềm tin của các cá nhân, gia đình hoặc nhóm chính trị bằng cách tạo nội dung giả mạo . Khi người ta không thể phân biệt được sự thật với hư cấu, niềm tin vào thông tin sẽ yếu đi, dẫn đến hậu quả nghiêm trọng trong việc xử lý tin giả, thông tin sai lệch.

- Sự gián đoạn chính trị và xã hội: Deepfake có thể được sử dụng để tạo ra hình ảnh sai lệch về các nhà lãnh đạo chính trị hoặc người nổi tiếng, gây rối và phá hoại bộ máy chính trị. Điều này có thể gây ra sự hoang mang và hỗn loạn trong xã hội, đặc biệt là trong các cuộc bầu cử hoặc những thời điểm quan trọng khác.

- Tăng rủi ro cho các cá nhân và tổ chức: Deepfakes có thể được sử dụng để đe dọa, phỉ báng hoặc gây hại cho các cá nhân hoặc tổ chức bằng cách tạo nội dung giả mạo, chẳng hạn như video khiêu khích hoặc nội dung sai sự thật, phỉ báng hoặc bôi nhọ.

- Gây hậu quả về tinh thần và cảm xúc: Nếu một người hoặc một tổ chức trở thành nạn nhân của giả mạo sâu, điều đó có thể gây ra những hậu quả nghiêm trọng về tâm lý và cảm xúc, chẳng hạn như căng thẳng, lo lắng và cảm thấy tổn thương lòng tự trọng.

# 

# VÍ DỤ THỰC TẾ VỀ VIỆC SỬ DỤNG DEEPFAKE ĐỂ GIẢ MẠO

## Quy trình để tạo ra Deepfake giả mạo khuôn mặt người khác

### Thu thập dữ liệu

**- Thu thập dữ liệu của người cần giả mạo (Người A)**:

+Tìm và tải xuống các hình ảnh và video của người A từ nhiều nguồn như internet, mạng xã hội, hoặc các video cá nhân. Đảm bảo có nhiều góc độ và biểu cảm khác nhau.

+Chuyển tất cả hình ảnh và video vào một thư mục riêng để tiện xử lý sau này.

**- Thu thập video của người cần thay thế (Người B)**:

+Chọn video mà bạn muốn thay thế khuôn mặt của người B bằng khuôn mặt của người A. Đảm bảo video này có chất lượng tốt và khuôn mặt của người B rõ ràng.

+Lưu video này vào một thư mục riêng biệt.

A collage of a person in a red suit

Description automatically generated

Hình 5‑1

### Chuẩn bị dữ liệu

**- Gắn thẻ và phát hiện khuôn mặt**:

**+Công cụ**: Sử dụng các thư viện như Dlib, OpenCV hoặc MTCNN để phát hiện và gắn thẻ khuôn mặt.

**+Quy trình**:

* Tải lên các hình ảnh và video của người A và người B.
* Chạy các công cụ phát hiện khuôn mặt để tự động nhận diện và đánh dấu các khuôn mặt trong từng khung hình.
* Lưu các khung hình đã được đánh dấu vào các thư mục riêng biệt cho người A và người B.

**+Làm sạch dữ liệu**:

* Kiểm tra và loại bỏ các khung hình không đạt yêu cầu (ví dụ: khuôn mặt bị che khuất, mờ hoặc không rõ ràng).
* Đảm bảo rằng các khung hình còn lại có chất lượng cao và rõ ràng.

A collage of a person

Description automatically generated

Hình 5‑2

### Huấn luyện mô hình

**- Chọn phần mềm và mô hình deepfake**:

**+Công cụ**: Tải xuống và cài đặt một phần mềm deepfake như DeepFaceLab, Faceswap, hoặc FakeApp.

**- Mô hình**: Sử dụng mô hình Autoencoders hoặc Generative Adversarial Networks (GANs).

**- Chuẩn bị môi trường huấn luyện**:

**+Phần cứng**: Sử dụng một máy tính với GPU mạnh mẽ (NVIDIA CUDA GPU được khuyến nghị).

**+Phần mềm**: Cài đặt các thư viện cần thiết như TensorFlow hoặc PyTorch tùy thuộc vào phần mềm bạn sử dụng.

**- Huấn luyện mô hình**:

**+Nạp dữ liệu**: Nạp các hình ảnh và video của người A và người B vào phần mềm.

**+Cấu hình mô hình**: Thiết lập các tham số huấn luyện như số lượng epoch, kích thước batch, và tỷ lệ học.

**+Bắt đầu huấn luyện**: Khởi động quá trình huấn luyện mô hình. Quá trình này có thể mất từ vài giờ đến vài ngày, tùy thuộc vào cấu hình phần cứng và khối lượng dữ liệu.

**+Giám sát**: Theo dõi tiến trình huấn luyện để đảm bảo mô hình học tốt. Điều chỉnh các tham số nếu cần thiết.

A screenshot of a video call

Description automatically generated

Hình 5‑3

### Tạo deepfake

**- Áp dụng mô hình đã huấn luyện**:

**+Quy trình**: Sau khi mô hình đã được huấn luyện, sử dụng nó để thay thế khuôn mặt của người B bằng khuôn mặt của người A trong video.

**+Công cụ**: Sử dụng các công cụ tích hợp trong phần mềm deepfake để thực hiện việc này.

**- Chỉnh sửa và tinh chỉnh**:

**+Công cụ**: Sử dụng các phần mềm chỉnh sửa video như Adobe Premiere Pro hoặc Final Cut Pro để tinh chỉnh các khung hình.

**+Quy trình**: Điều chỉnh các khung hình để đảm bảo tính chân thực và liên tục của video. Loại bỏ các khung hình có lỗi hoặc không khớp với khuôn mặt giả mạo.

**+Hiệu ứng**: Áp dụng các hiệu ứng hậu kỳ để làm mịn và cải thiện chất lượng video.

### Kiểm tra và hoàn thiện

**- Kiểm tra video deepfake**:

**+Quy trình**: Xem lại toàn bộ video để phát hiện bất kỳ lỗi nào. Đảm bảo rằng khuôn mặt giả mạo trông tự nhiên và khớp với các phần còn lại của video.

**+Chỉnh sửa bổ sung**: Thực hiện các chỉnh sửa bổ sung nếu cần thiết.

**- Áp dụng hiệu ứng hậu kỳ**:

**+Công cụ**: Sử dụng phần mềm chỉnh sửa video để thêm các hiệu ứng hậu kỳ như điều chỉnh màu sắc, ánh sáng, và các yếu tố khác.

**+ Quy trình**: Làm mịn và cải thiện chất lượng video để làm cho video trông chân thực hơn.

A collage of a person holding up his hand

Description automatically generated

Hình 5‑4

## Công cụ và phần mềm phát triển

**- DeepFaceLab**: Một trong những công cụ mạnh mẽ và phổ biến nhất để tạo deepfake. Hỗ trợ nhiều tùy chọn và cấu hình khác nhau.

**- Faceswap**: Công cụ mã nguồn mở được sử dụng rộng rãi, cung cấp nhiều tùy chọn huấn luyện và ứng dụng.

**- FakeApp**: Dễ sử dụng nhưng ít mạnh mẽ hơn so với DeepFaceLab và Faceswap.

* Video chi tiết tự deepfake mặt bản thân: video 1 -[link](https://drive.google.com/file/d/1ClpFZCI1j5JhTl2abulSMUktoRIi9W0J/view?usp=sharing)

Video 2 – [link](https://drive.google.com/file/d/1ln4rs66a2LXjxYx_TnYISmtY-vTfqkru/view?usp=sharing)

# CÁC BIỆN PHÁP PHÒNG NGỪA VÀ GIẢI QUYẾT

- Những gì công nghệ deepfake có thể làm là khiến nhiều người phát điên khi nghĩ rằng họ đang bị lợi dụng. Một số vụ bê bối nghiêm trọng liên quan đến an ninh của các cá nhân và cộng đồng có liên quan đến Deepfake đã xảy ra trên toàn thế giới, không chỉ ảnh hưởng đến những người nổi tiếng và quyền lực mà còn cả những người bình thường và các chuyên gia, những người đã phơi bày mối nguy hiểm mà Deepfake thực sự có thể gây ra. Nhiều nghiên cứu đang được tiến hành về các giải pháp ngăn chặn phần mềm độc hại deepfake.

- Trước hết, để không trở thành nạn nhân của Deepfake, mọi người cần đảm bảo thông tin mình chia sẻ có phải là thông tin gốc hay không, đồng thời nhận diện các dấu hiệu hoạt động tội phạm mạng theo các yếu tố sau:

- Chuyển động trong video giật lag, như đoạn video hỏng; ánh sáng liên tục thay đổi từ khung hình này khung hình khác; thay đổi liên tục về tình trạng da; một vài sự rung chuyển bất thường được nhìn thấy trong video.

- Khẩu hình miệng và lời nói không khớp; chất lượng âm thanh hoặc video thấp; nhân vật nói liên tục không chớp mắt.

- Trong trường hợp nhận được cuộc gọi video có dấu hiệu trên, người dân cần bình tĩnh và tiến hành xác minh thông tin trực tiếp với người gọi qua kênh khác như gọi di động, nhắn tin xác nhận... Trường hợp phát hiện video giả mạo, nhạy cảm hoặc có chứa nội dung tuyên truyền các quan điểm sai trái, thù địch, cần liên hệ với nhà cung cấp dịch vụ để yêu cầu các cơ quan này hỗ trợ trong việc xóa bỏ, khóa hoặc gỡ bỏ các nội dung Deepfake có hại cho mình và xã hội.

- Bạn nên cẩn thận trước khi nhấp vào bất kỳ liên kết bất thường nào xuất hiện trên mạng xã hội hoặc liên kết chứa mã độc; đảm bảo rằng hệ thống bảo mật của phần mềm luôn được cập nhật,hãy chủ động bảo vệ thông tin cá nhân cẩn thận và cập nhật phần mềm thường xuyên.

- Khi truy cập Internet công cộng, bạn phải tự nguyện đăng xuất khỏi tài khoản trang sau khi sử dụng, thường xuyên thay đổi mật khẩu và sử dụng mật khẩu khó đoán để tăng thêm tính bảo mật. Hạn chế chia sẻ thông tin cá nhân như hình ảnh, video, âm thanh của bạn hoặc người thân với những người mà bạn không biết hoặc không tin tưởng.

- Các nhà quản lý cần cải thiện khả năng công nghệ của mình để chống lại các thủ đoạn tinh vi mới xuất hiện.Trong bối cảnh đất nước đang đứng trước nhiều thay đổi to lớn, trước những âm mưu, thủ đoạn chống phá ngày càng điên cuồng của các thế lực thù địch và nhiều thủ đoạn hoạt động xảo quyệt của các loại tội phạm - trong đó không loại trừ việc sử dụng Deepfake như một thứ “vũ khí” mới, các cơ quan chức năng Việt Nam cần tăng cường công tác bảo vệ chính trị nội bộ, huy động được sức mạnh của toàn dân, phối hợp cùng nhau trong tham mưu cho Đảng, Nhà nước xây dựng biện pháp đối phó, phối hợp phòng ngừa, phát hiện, ngăn chặn đấu tranh xử lý hiệu quả những hành vi lợi dụng công nghệ mới để phá hoại an ninh quốc gia, trật tự an toàn xã hội, giữ vững sự ổn định và phát triển của đất nước. [5]



# PHÂN TÍCH

## Phân tích các nghiên cứu hiện có và thảo luận trong cộng đồng học thuật và chính trị về ảnh hưởng của Deepfake:

Các nghiên cứu hiện có đã chỉ ra nhiều khía cạnh khác nhau về ảnh hưởng của Deepfake. Một số nghiên cứu tập trung vào việc phát triển các công cụ và phương pháp để phát hiện Deepfake, như sử dụng các thuật toán học máy để nhận diện những bất thường trong video hoặc âm thanh. Các nghiên cứu khác lại tập trung vào việc đánh giá tác động xã hội của Deepfake, bao gồm khả năng làm giảm sự tin tưởng vào truyền thông và gây ra những hậu quả nghiêm trọng trong các lĩnh vực như chính trị, pháp luật và an ninh.

Trong cộng đồng học thuật, các thảo luận thường xoay quanh việc làm thế nào để cân bằng giữa sự tiến bộ của công nghệ và những rủi ro đạo đức mà nó mang lại. Các nhà nghiên cứu nhấn mạnh tầm quan trọng của việc phát triển các biện pháp kỹ thuật để phát hiện Deepfake, cũng như việc xây dựng các khung pháp lý và chính sách quản lý phù hợp.

Trên mặt trận chính trị, Deepfake đã được nhận diện như một mối đe dọa tiềm tàng đối với an ninh quốc gia và sự ổn định xã hội. Nhiều quốc gia đang xem xét hoặc đã triển khai các quy định pháp luật nhằm kiểm soát việc tạo ra và phân phối Deepfake, nhằm bảo vệ công dân khỏi các thông tin giả mạo và giảm thiểu các nguy cơ liên quan.

- Trong nghiên cứu khoa học xã hội và nhân văn về deepfake, các nghiên cứu điện ảnh cho thấy rằng khuôn mặt con người trở thành trọng tâm của sự xung đột trong thời đại số. Deepfake đã được sử dụng để tái hiện lại lịch sử điện ảnh bằng cách đưa các diễn viên mới vào các bộ phim kinh điển. Ngoài các vấn đề như nội dung khiêu dâm, deepfake còn đối mặt với các mối đe dọa về nhận thức và xã hội, gây ra những hệ lụy nghiêm trọng như bóc lột, đe dọa và phá hoại cá nhân. Các cuộc thảo luận học thuật cũng tập trung vào các phản ứng pháp lý và quy định tiềm năng, đặc biệt trong lĩnh vực nghiên cứu pháp lý và truyền thông, để đối phó với những tác hại của deepfake như thông tin sai lệch và tác động xã hội.

- Nghiên cứu khoa học máy tính về deepfake đã có những bước tiến đáng chú ý. Chương trình Video Rewrite, ra mắt năm 1997, là dự án đầu tiên tự động tái tạo khuôn mặt từ âm thanh khác. Chương trình "Tổng hợp Obama" (2017) cải tiến kỹ thuật bằng cách đồng bộ hóa hình ảnh miệng với âm thanh. Face2Face (2016) là phương pháp đầu tiên tái hiện nét mặt thời gian thực mà không cần máy ảnh chuyên nghiệp. Đại học California, Berkeley, vào năm 2018, giới thiệu ứng dụng khiêu vũ giả, mở rộng deepfake vào lĩnh vực toàn thân và giả mạo hình ảnh y tế, thậm chí tấn công vào bệnh viện trong thử nghiệm thâm nhập mũ trắng.

[4] Hầu hết các tác phẩm deepfake trên các nền tảng xã hội như YouTube hay Facebook ngày nay có thể được coi là những tác phẩm nghệ thuật hoặc trò vui vô hại sử dụng các nhân vật công chúng, dù còn sống hay đã chết. Tuy nhiên, deepfake cũng có mặt tối, bao gồm nội dung khiêu dâm về người nổi tiếng và trả thù, cũng như những nỗ lực nhằm gây ảnh hưởng chính trị và phi chính trị.

Nhiều tác phẩm deepfake tập trung vào những người nổi tiếng, chính trị gia và lãnh đạo doanh nghiệp vì Internet chứa đầy ảnh và video về họ, cung cấp kho lưu trữ hình ảnh lớn cần thiết để đào tạo hệ thống AI. Phần lớn các tác phẩm này là những trò đùa và meme hài hước. Tuy nhiên, các ví dụ về deepfake có hại ngày càng xuất hiện nhiều hơn.

Công nghệ deepfake cho phép tạo ra nội dung khiêu dâm không tự nguyện sử dụng hình ảnh của người nổi tiếng và người bình thường, chia sẻ trên mạng xã hội mà không có sự đồng ý của họ. Ngoài ra, deepfake còn được sử dụng để thao túng dữ liệu và tạo ra các video nhằm gây hiểu lầm hoặc ảnh hưởng đến nhận thức của công chúng.

[5] Gần đây, những hình ảnh bịa đặt lan truyền trên mạng về cuộc bầu cử tổng thống Mỹ, bao gồm hình ảnh cựu Tổng thống Donald Trump được bao quanh bởi các nhóm người da đen đang mỉm cười và nô đùa, như thể ông đang thu phục cử tri da đen. Các cuộc thăm dò cho thấy cử tri da đen vẫn trung thành với Tổng thống Joe Biden. Những hình ảnh này củng cố cảnh báo rằng việc sử dụng hình ảnh do AI tạo ra sẽ gia tăng khi cuộc tổng tuyển cử tháng 11 đến gần.

[6] Ấn Độ, tháng 4 năm 2018: Một video lan truyền trên WhatsApp cho thấy một nhóm trẻ em chơi cricket trên đường phố, rồi hai người đàn ông trên xe máy bắt cóc một đứa trẻ. Video này gây hoang mang và bạo lực, khiến ít nhất chín người vô tội thiệt mạng. Thực tế, video là bản chỉnh sửa từ một chiến dịch giáo dục ở Pakistan về nhận thức bắt cóc trẻ em, phần cảnh báo đã bị cắt bỏ.

Cùng tháng, BuzzFeed đăng video cựu Tổng thống Obama nói trong một phòng Oval Office giả tạo, với 35 giây đầu ông đưa ra tuyên bố sốc: "Tổng thống Trump là một tên ngu ngốc hoàn toàn và toàn diện," trước khi tiết lộ rằng video là giả.Gần đây, thuật ngữ “deepfake” được đề cập tới nhiều, đặc biệt xung quanh việc phát tán ảnh khỏa thân giả của ca sĩ nổi tiếng Taylor Swift trên X. [7] Sự gia tăng sử dụng AI mang lại cả lợi ích và những vấn đề mới. Deepfake là hình ảnh hoặc video do AI tạo ra, rất chân thực và khó phân biệt với nội dung thật. Kỹ thuật này có thể thay thế và kết hợp khuôn mặt một người vào hình ảnh hoặc video khác, dẫn đến việc tạo ra các nội dung giả mạo, như hình ảnh khiêu dâm của Taylor Swift bị ghép từ ảnh của cô.

Deepfake lan truyền nhanh chóng, làm khó khăn việc xác thực thông tin trên mạng xã hội. Đội ngũ truyền thông của người nổi tiếng và chính trị gia cần có nguồn lực để theo dõi và phản ứng kịp thời với tin giả. Việc xác thực thông tin chính thống trở nên phức tạp hơn, đòi hỏi kỹ thuật kiểm chứng chuyên sâu.

Trong bối cảnh deepfake lan tràn, chiến lược quản lý rủi ro hiệu quả nhất là duy trì các kênh thông tin nhất quán thông qua mạng xã hội, website, và gặp mặt trực tiếp giữa doanh nghiệp, người nổi tiếng, chính trị gia với các bên liên quan như người hâm mộ, báo chí, và cộng đồng.So sánh với các công nghệ và hiện tượng tương tự khác như tin tức giả mạo, sự thay đổi hình ảnh,…

## So sánh với các công nghệ và hiện tượng tương tự khác như tin tức giả mạo, sự thay đổi hình ảnh,…

[4] Trong những năm gần đây, tin tức giả đã trở thành một vấn đề đe dọa đến diễn ngôn công chúng, xã hội và nền dân chủ. Tin giả là nội dung hư cấu được bịa đặt để đánh lừa công chúng, lan truyền nhanh chóng trên mạng xã hội và có thể ảnh hưởng đến hàng triệu người dùng. Hiện tại, cứ năm người dùng Internet thì có một người nhận được tin tức qua YouTube, chỉ sau Facebook.

Sự phổ biến của video làm nổi bật nhu cầu xác minh tính xác thực của nội dung, vì công nghệ mới cho phép thao tác video một cách thuyết phục. Deepfakes, video siêu thực tế bằng cách hoán đổi khuôn mặt, được tạo ra bởi AI và không để lại dấu vết thao túng. Điều đáng lo ngại là phạm vi, quy mô và độ phức tạp của công nghệ deepfake, vì bất kỳ ai có máy tính đều có thể tạo ra video giả khó phân biệt với video thật.• Deepfake có những điểm tương đồng với các công nghệ và hiện tượng như tin tức giả mạo (fake news) và sự thay đổi hình ảnh (image manipulation). Tuy nhiên, Deepfake vượt trội hơn ở mức độ chân thực và khó phát hiện hơn. Trong khi các hình ảnh giả mạo truyền thống có thể được phát hiện qua các dấu hiệu chỉnh sửa rõ ràng, Deepfake sử dụng AI để tạo ra các nội dung gần như hoàn hảo, làm cho việc phân biệt trở nên khó khăn hơn nhiều.

• So với tin tức giả mạo, Deepfake có thể tác động mạnh mẽ hơn vì khả năng kết hợp hình ảnh và âm thanh để tạo ra những nội dung giả mạo rất thuyết phục. Điều này có thể làm tăng nguy cơ của thông tin sai lệch lan rộng và gây ra hậu quả nghiêm trọng hơn đối với xã hội.

[6] Trong khi việc "photoshopping" ảnh tĩnh đã từ lâu trở thành một phần không thể thiếu trong văn hóa số, video bị chỉnh sửa của con người ngày càng xuất hiện nhiều hơn trên mạng. Dựa trên thuật toán học máy, các ứng dụng phần mềm có thể tạo ra những video "ghép mặt" rất thuyết phục, trong đó biểu cảm của một người được đặt cẩn thận lên đầu của người khác. Các ứng dụng này cũng có thể sử dụng bản ghi âm của chuyển động miệng và giọng nói của một người để phục hồi lại lời nói của họ và khiến họ nói bất kỳ câu nào. Kết quả có thể thuyết phục đáng kinh ngạc, đặc biệt là với những video có độ phân giải thấp phổ biến trên mạng. Hình ảnh có sức thuyết phục mạnh mẽ hơn văn bản và người xem thường có sự phòng vệ kém hơn đối với sự lừa dối hình ảnh như vậy.

# KẾT LUẬN VÀ ĐỀ XUẤT

## Tóm tắt nhấn mạnh tầm quan trọng của vấn đề

Qua phân tích, chúng ta thấy rằng Deepfake là một công nghệ có tiềm năng to lớn nhưng cũng đặt ra nhiều thách thức nghiêm trọng. Từ việc gây nhiễu loạn thông tin, làm giảm lòng tin của công chúng vào truyền thông, đến những nguy cơ đối với an ninh quốc gia và sự ổn định xã hội, Deepfake đòi hỏi một sự quan tâm và quản lý cẩn thận từ cả cộng đồng học thuật, chính trị và xã hội.

- Chính trị: Deepfake có thể được sử dụng để tạo ra video giả mạo của các nhà lãnh đạo chính trị, gây ảnh hưởng đến quá trình bầu cử hoặc tạo ra sự hoang mang trong dư luận. Các video Deepfake có thể lan truyền thông tin sai lệch và ảnh hưởng đến sự tin tưởng của công chúng đối với chính trị gia.

- Quảng cáo: Deepfake có thể được sử dụng để tạo ra quảng cáo gian lận hoặc đánh lừa người tiêu dùng. Điều này có thể gây ảnh hưởng tiêu cực đến niềm tin và quyết định mua hàng của người tiêu dùng.

- Giải trí: Deepfake đã tạo ra sự thay đổi lớn trong ngành công nghiệp giải trí. Nó có thể được sử dụng để tạo ra video giả mạo của các diễn viên nổi tiếng hoặc tái hiện lại những cảnh quay nổi tiếng. Điều này có thể gây nhầm lẫn và tranh cãi về tính chân thực của nội dung giải trí.

- Quan hệ cá nhân: Deepfake có thể tạo ra video giả mạo của người khác và lan truyền thông tin sai lệch, gây hại đến đời sống riêng tư và tạo ra xung đột trong quan hệ cá nhân.

## Đề xuất các biện pháp cụ thể để ứng phó với ảnh hưởng của Deepfake đối với xã hội và đạo đức

• Phát triển các công nghệ phát hiện Deepfake: Đầu tư vào nghiên cứu và phát triển các công cụ kỹ thuật để phát hiện Deepfake một cách hiệu quả và nhanh chóng.

• Thiết lập các khung pháp lý và chính sách: Xây dựng và thực thi các quy định pháp luật nghiêm ngặt về việc tạo ra, phân phối và sử dụng Deepfake, bao gồm việc xử phạt các hành vi vi phạm.

• Tăng cường giáo dục và nhận thức cộng đồng: Tổ chức các chiến dịch giáo dục để nâng cao nhận thức của công chúng về nguy cơ của Deepfake, giúp họ nhận biết và phòng tránh các nội dung giả mạo.

• Khuyến khích sự hợp tác quốc tế: Hợp tác với các quốc gia khác trong việc chia sẻ thông tin, công nghệ và kinh nghiệm để đối phó với Deepfake ở quy mô toàn cầu.

• Nghiên cứu liên ngành: Thúc đẩy các nghiên cứu liên ngành giữa công nghệ, luật pháp, xã hội học và đạo đức học để có cách tiếp cận toàn diện hơn trong việc giải quyết các thách thức do Deepfake mang lại.

[8] - Tính chất vừa mới mẻ vừa phức tạp của công nghệ deepfake khiến việc truy nguồn tạo và phát tán các nội dung sai lệch này rất khó. Giới chuyên gia cho rằng các chính phủ và công ty công nghệ chưa đủ khả năng ngăn ngừa vấn nạn này, thậm chí còn chưa hành động đủ nhanh và mạnh mẽ để ứng phó. Tháng 9-2023, dự luật đề xuất cấm sử dụng nội dung lừa dối do AI tạo làm ảnh hưởng tới bầu cử đã được trình lên Quốc hội Mỹ. Trong khi chờ đợi, một số bang đã có các quy định riêng để đối phó với nạn deepfake.

- Gần đây nhất, Nghị viện châu Âu đã phê chuẩn Đạo luật AI, cấm việc dùng AI tạo hình ảnh gương mặt từ nguồn là các video giám sát. EU cũng kêu gọi 8 công ty công nghệ lớn - trong đó có Google, Facebook và X - cung cấp thông tin chi tiết về cách các hãng này phát hiện và xử lý nội dung deepfake. Báo The Guardian bình luận rằng EU có thể vận dụng đạo luật mới để yêu cầu các hãng công nghệ phải loại bỏ tận gốc hình ảnh, video và âm thanh giả mạo.

- Anh cũng vừa thông qua Đạo luật an toàn trực tuyến, tiến tới quản lý một số khía cạnh của AI, chẳng hạn như hình sự hóa hành vi chia sẻ deepfake mà không có sự đồng thuận của đối tượng liên quan.

[9] Trên thế giới, nhiều cơ chế và chính sách đã được ban hành yêu cầu doanh nghiệp vận dụng các giải pháp an ninh mạng. Cụ thể, giới chức tại Mỹ đã ban hành các chính sách quản lý việc thu thập và xử lý dữ liệu tài chính và thẻ tín dụng, cùng với các chính sách an ninh mạng cho dữ liệu lưu trữ trên điện toán đám mây. Tại khu vực Đông Nam Á, Hiệp hội các quốc gia Đông Nam Á (ASEAN) đã thống nhất đề ra Chiến lược Hợp tác an ninh mạng, bao gồm bảo vệ dữ liệu cá nhân, bảo mật trong lưu trữ dữ liệu, và bảo đảm quyền lợi khách hàng liên quan đến an toàn thông tin.

[5] Không chỉ Mỹ, EU cũng sắp thông qua một trong những luật toàn diện nhất để quản lý AI, nhưng luật này phải đến năm 2026 mới có hiệu lực. Quy định được đề xuất ở Anh cũng đã bị chỉ trích vì tiến triển quá chậm.

\*Phần lớn các chuyên gia về bảo mật đều đưa ra những lời khuyên nhằm phòng tránh trước các tính huống xấu có thể xảy ra, đặc biệt là việc nâng cao bảo mật các tài khoản trực tuyến.

- Tin tưởng nhưng vẫn phải xác minh: Nếu nhận được một đoạn tin nhắn thoại đặc biệt là các tin nhắn hướng đến mục đích vay tiền hoặc mượn tài sản cá nhân của bạn. Dù cho giọng nói nghe rất quen thuộc và cực kỳ giống thật, tốt nhất bạn vẫn nên gọi lại bằng số điện thoại mà bạn biết là chính xác để xác minh rằng đồng nghiệp, người thân đó của bạn đã thực sự gửi yêu cầu bằng tin nhắn thoại.

- Đừng vội truy cập bất kỳ đường link nào: Khi được người thân gửi cho một đường link với những cấu trúc lạ, đừng vội vàng nhấp vào đường link/nút đó. Rất có thể đó là một cái bẫy được đưa ra để kẻ lừa đảo có thể chiếm đoạt tài khoản trực tuyến của bạn. Hãy tìm cách xác minh lại với người đã gửi tin nhắn cho bạn.

- Hãy chú ý vào những điều nhỏ nhất: Nếu nhận được một cuộc gọi chuyển tiền hoặc video call với nội dung chuyển tiền, dù là người thân thiết, bạn vẫn hãy xem kỹ số điện thoại, email hoặc tài khoản có trùng hợp với người đang thực hiện yêu cầu hay không. Thông thường các đối tượng sẽ yêu cầu chuyển tiền sang một tài khoản bên thứ ba hoặc sang một tài khoản có tên gần tương tự.

- Hạn chế quyền truy cập vào giọng nói và hình ảnh của bạn: Như chúng ta đã biết, để có thể tạo ra hình ảnh và âm thanh giả những kẻ lừa đảo sẽ cần các bản ghi âm, hình ảnh hoặc cảnh quay của bạn để tạo ra những sản phẩm giả. Để ngăn chặn việc dữ liệu hình ảnh, âm thanh của bạn bị sao chép, hãy hạn chế sự hiện diện của bạn trên mạng xã hội hoặc đặt tài khoản của bạn ở chế độ riêng tư, chỉ chấp nhận yêu cầu từ những người bạn tin tưởng.

- Không những vậy, một số các chuyên trang về bảo mật còn đưa ra lời khuyên để có thể phát hiện deepfake ngay tức thì như sau:

+ Chuyển động giật cục, như một đoạn video lỗi

+ Ánh sáng bị thay đổi liên tục từ khung hình này sang khung hình tiếp theo

+ Thay đổi tông màu da liên tục

+ Video có những sự nhấp nháy lạ thường

+ Khẩu hình miệng không đồng bộ với lời nói

+ Hiện lên các đồ vật kỹ thuật số trong hình ảnh

+ Âm thanh và/hoặc video chất lượng thấp

+ Nhân vật nói liên tục, không chớp mắt

[10] Ngày 12/06/2018, Quốc hội Việt Nam đã thông qua Luật An ninh mạng, có hiệu lực từ ngày 01/01/2019, quy định về hoạt động bảo vệ an ninh quốc gia và bảo đảm trật tự, an toàn xã hội trên không gian mạng; trách nhiệm của cơ quan, tổ chức, cá nhân có liên quan. Đáng chú ý, có 6 nhóm hành vi bị nghiêm cấm thực hiện trên không gian mạng, đặc biệt là việc sử dụng không gian mạng, công nghệ thông tin, phương tiện điện tử vi phạm an ninh quốc gia, trật tự an toàn xã hội; các hoạt động chống Nhà nước; xuyên tạc lịch sử, xúc phạm tôn giáo, phân biệt đối xử về giới, phân biệt chủng tộc; kích động gây bạo loạn, phá rối an ninh, gây rối trật tự công cộng. Thông tin sai sự thật, làm nhục, vu khống; xâm phạm trật tự quản lý kinh tế, hoạt động mại dâm, tệ nạn xã hội, mua bán người và xúi giục, lôi kéo, kích động người khác phạm tội.

# CÁC VÍ DỤ VỀ LỢI ÍCH CỦA DEEPFAKE TỪ TRƯỚC TỚI NAY.

## Bùng nổ nhu cầu dùng deepfake nói chuyện với người đã khuất

- Phần mềm chỉnh sửa ảnh giúp "đoàn tụ" những người mất tích - MyHeritage.

- Mặc dù có vẻ như không thể nhìn thấy ngay một người thân đã khuất đang làm việc và mỉm cười trong một đoạn video mà bạn chưa từng quay, nhưng điều này giờ đây đã trở thành hiện thực với công nghệ đang phát triển. Deepfake là một công nghệ đã có mặt trên thị trường được vài năm và chủ yếu dựa trên "mạng thần kinh nhân tạo", tức là hệ thống máy tính có thể nhận dạng mẫu trong dữ liệu. Có rất nhiều ứng dụng và trang web cung cấp các công cụ toàn diện và tính năng hoán đổi khuôn mặt ngày càng trở nên phổ biến trên TikTok. Việc nâng cơ mặt này sử dụng công nghệ tiên tiến cũng khiến nhiều người giống như những ngôi sao, và sự thật là những người này giống nhau đến mức không thể phân biệt được với nhau.

- Nhưng phương pháp nghiên cứu phả hệ trực tuyến mới MyHeritage hơi khác một chút, nó cũng sử dụng công nghệ sâu, nhưng không phải ở hình thức bên ngoài. Với ứng dụng này, người dùng chỉ cần tải ảnh của người thân đã khuất của mình lên và ứng dụng sẽ biến bức ảnh đó thành ảnh động hoặc video để người đã khuất có thể mỉm cười và cử động như người thật. MyHeritage là trang phổ biến và được nhiều người yêu thích nên ít có nguy cơ hình ảnh/dữ liệu của người dùng bị sử dụng sai mục đích như nhiều ứng dụng khác. Tuy nhiên, cũng có nhiều ý kiến ​​trái chiều xuất hiện. Số lượng người sử dụng các ứng dụng “hẹn hò” hiện nay rất cao và nhiều người trong số họ trở nên rất xúc động khi “gặp nhau”. Một cư dân mạng cho biết cô đã khóc rất nhiều khi được “đoàn tụ” với người mẹ quá cố. Một người khác đã dùng ứng dụng này để làm ông nội sống lại rồi đưa ông ra mắt gia đình và cả gia đình cùng khóc.

- Nhưng cũng có người dùng trang này nói rằng kết quả của chương trình này có vẻ là thật, chân thực đến mức khiến một số người "sợ hãi" hoặc "ám ảnh" và rằng bức ảnh của người thân đã khuất là " người thực". ", di chuyển, nháy mắt, mỉm cười... Gilad Japhet, người sáng lập và CEO của MyHeritage, cho biết mục đích "gặp gỡ người thân đã khuất" là giúp mọi người "liên lạc với lịch sử gia đình" và đặc biệt hữu ích cho những gia đình không có gia đình. Không có bức ảnh nào của người thân đã khuất.

## Deepfake triển khai cho digital marketing

Không cần sử dụng diễn viên trực tiếp trong video hoạt hình. Ngày nay, deepfake có thể là một giải pháp mới khi công ty chỉ cần mua giấy phép sử dụng hình ảnh của người chơi, tạo hoặc sử dụng lại video quảng cáo cũng như thêm hình ảnh và lời thoại phù hợp để tạo ra sản phẩm quảng cáo mới. Thậm chí, công ty còn có thể tiết kiệm được số tiền khổng lồ nếu sử dụng nhân sự nội bộ để tạo video. Tóm lại, Deepfake sẽ rất hữu ích cho những nhà tiếp thị có ngân sách eo hẹp nhưng muốn thu được kết quả tốt nhất từ ​​quảng cáo của mình.

A group of men standing behind a device

Description automatically generated

Hình 9‑1

## Deepfake ứng dụng trong giáo dục:

Năm 2017, Đại học Washington đã sử dụng deepfake để giới thiệu Abraham Lincoln phát biểu tại Gettysburg, trong khi đó, Tạp chí TIME kết hợp cùng với Đại học Stanford đã sử dụng deepfake để chiếu cảnh Martin Luther King Jr. đọc bài phát biểu “Tôi có một giấc mơ” ở một sự kiện trong năm 2020.

## Deepfake dùng trong làm film

Chắc bạn vẫn còn nhớ nhà sản xuất Fast and Furious 7 đã phải sử dụng Deepfake lên chính người em của cố nam diễn viên Paul Walker để giúp cho nhân vật Brian O'Conner có thể hoàn thành cái kết thật trọn vẹn đúng không nào?



Hình 9‑2

# 

# CÁC VÍ DỤ VỀ HẬU QUẢ CỦA DEEPFAKE TỪ TRƯỚC TỚI NAY

## Video tổng thống obama được tạo bởi deepfake

Vào ngày 17 tháng 4 năm 2018, một đoạn video cho thấy sự giống nhau của cựu Tổng thống Mỹ Barack Obama đã xuất hiện trên mạng. Trong video này, một nhân vật giống ông Obama dùng những từ ngữ xúc phạm để chỉ trích đương kim Tổng thống Mỹ Donald Trump. Dù sau đó dư luận mới phát hiện ra đoạn video đó là giả nhưng nó lại lan truyền nhanh chóng và ít nhiều dẫn đến những hậu quả không đáng có.A person in a suit

Description automatically generated

Hình 10‑1

## Các sản phẩm đồi trụ có gương mặt người nổi tiếng

- Trong suốt một tháng qua, tài khoản Twitter liên tục đăng tải những bức ảnh vui vẻ của các sao nam. Tài khoản thông báo: "Chúng tôi chấp nhận ảnh giả mạo của các nghệ sĩ, TikTokers và lượt thích của bạn theo yêu cầu. Chính sách quyền riêng tư (phí hộp thư đến). Tất cả chỉ nhằm mục đích giải trí, phi tội phạm." Nhiều ca sĩ, diễn viên hay gương mặt nổi tiếng trên trang Social Media tại Việt Nam đã trở thành nạn nhân như Sơn Tùng M-TP, HIEUTHUHAI, Quân A.P, Liên Bỉnh Phát... và người cuối cùng là Ngô Kiến Huy. Khuôn mặt của họ được đóng khung với vẻ ngoài đáng xấu hổ và khiêm tốn. - Việc tạo video bằng công nghệ Deepfake đang trở nên dễ dàng hơn bao giờ hết, theo Forbes. Phần mềm dùng để tạo ra những video giả mạo này được phát tán rộng rãi trên các trang chia sẻ như Github hay các diễn đàn internet phổ biến như 4chan. Giờ đây, bất kỳ người dùng nào có kiến ​​thức kỹ thuật và phần cứng mạnh mẽ đều có thể tải xuống công cụ và tự tạo các video sống động.

Hình 10‑2

## Cuộc gọi video Deepfake để lừa đảo

Anh trai cô Tâm ở Mỹ đã liên hệ tài khoản của cô để vay tiền vì có việc gấp. Lúc này, cô Tâm nhận được video qua ứng dụng Messenger của Facebook. “Cuộc trò chuyện kéo dài vài giây nhưng cũng đủ để tôi nhìn thấy khuôn mặt người thân của mình và anh ta nhìn tôi bằng ánh mắt quen thuộc nhưng đáng sợ. Sau đó tôi nhận được thông báo rằng tôi không thể tiếp tục video vì tín hiệu yếu.  “Tôi chuyển sang nhắn tin để công việc dễ dàng hơn. Tôi cũng tin tưởng gửi tiền theo hướng dẫn”. Chị Tâm cho biết

Hình 10‑3

# KẾT LUẬN

Công nghệ giả mạo sâu (deepfake) mang đến những tiềm năng vượt bậc trong nhiều lĩnh vực, từ giải trí, quảng cáo, giáo dục đến y học. Nhờ khả năng tạo ra những nội dung số có độ chân thực cao, deepfake có thể thay đổi cách chúng ta tạo ra và tiêu thụ thông tin. Tuy nhiên, cùng với những tiềm năng này, deepfake cũng đặt ra những thách thức lớn đối với xã hội và đạo đức.

Trước hết, các tác động tiêu cực của deepfake, như việc lan truyền thông tin sai lệch và lừa đảo, có thể gây ra những hậu quả nghiêm trọng. Những video giả mạo có thể được sử dụng để bôi nhọ danh tiếng của các cá nhân hoặc tổ chức, gây hoang mang trong dư luận và thậm chí kích động xung đột xã hội. Ngoài ra, deepfake còn có thể được sử dụng trong các hoạt động tống tiền hoặc lừa đảo, gây tổn thất nặng nề về mặt tinh thần và tài chính cho các nạn nhân.

Sự khó khăn trong việc phân biệt giữa thật và giả do deepfake gây ra còn làm suy giảm niềm tin của công chúng vào truyền thông và các nguồn thông tin trực tuyến. Khi công chúng không còn tin tưởng vào những gì họ thấy và nghe, điều này có thể dẫn đến một trạng thái hoang mang và mất niềm tin rộng rãi, ảnh hưởng đến sự ổn định của xã hội và làm rạn nứt các mối quan hệ cá nhân.

Từ góc độ đạo đức, deepfake đặt ra nhiều câu hỏi phức tạp về quyền riêng tư và sự tôn trọng đối với cá nhân. Việc sử dụng hình ảnh và giọng nói của người khác mà không có sự đồng ý rõ ràng không chỉ xâm phạm quyền riêng tư mà còn vi phạm các nguyên tắc đạo đức cơ bản về sự tôn trọng và sự đồng ý cá nhân. Các quy định pháp lý hiện tại ở nhiều quốc gia vẫn chưa theo kịp sự phát triển nhanh chóng của công nghệ này, đòi hỏi các nhà làm luật phải nhanh chóng xây dựng và cập nhật các quy định để bảo vệ quyền lợi của công dân và duy trì trật tự xã hội.

Để đối phó với những thách thức mà deepfake mang lại, cần có sự hợp tác chặt chẽ giữa các nhà khoa học, các công ty công nghệ và các cơ quan quản lý. Nghiên cứu và phát triển các công cụ phát hiện và ngăn chặn deepfake là cần thiết hơn bao giờ hết. Các công nghệ nhận diện hình ảnh và giọng nói, kết hợp với trí tuệ nhân tạo, có thể giúp phát hiện các dấu hiệu bất thường trong các nội dung số, từ đó ngăn chặn sự phát tán của các nội dung giả mạo.

Đồng thời, việc xây dựng và cập nhật các quy định pháp lý để quản lý và kiểm soát deepfake là điều vô cùng quan trọng. Các cơ quan quản lý cần đưa ra các biện pháp pháp lý nghiêm ngặt để ngăn chặn việc lạm dụng công nghệ này, đồng thời bảo vệ quyền lợi của các cá nhân và tổ chức. Cùng với đó, việc giáo dục cộng đồng về cách nhận diện thông tin giả mạo và tăng cường nhận thức về đạo đức trong việc sử dụng công nghệ cũng là những bước quan trọng nhằm giảm thiểu tác động tiêu cực của deepfake.

Công nghệ deepfake, với những tiềm năng và thách thức của nó, đòi hỏi một cách tiếp cận cân bằng và có trách nhiệm. Chỉ khi chúng ta hiểu rõ và kiểm soát được công nghệ này, chúng ta mới có thể khai thác tối đa những lợi ích mà nó mang lại, đồng thời bảo vệ xã hội khỏi những hệ lụy tiêu cực. Hợp tác chặt chẽ giữa các bên liên quan và sự đồng lòng của toàn xã hội sẽ là chìa khóa để đối phó với các thách thức mà deepfake đặt ra, hướng tới một tương lai mà công nghệ được sử dụng một cách có trách nhiệm và đạo đức.

# Tài liệu tham khảo

|  |
| --- |
| 1. Hồng Minh, “Deepfake Những nguy cơ tiềm ẩn cho xã hội và cá nhân,” TRANG THÔNG TIN ĐIỆN TỬ THÀNH PHỐ BIÊN HÒA - ĐỒNG NAI, 12 6 2023. [Trực tuyến]. Available: https://bienhoa.dongnai.gov.vn/Pages/   newsdetail.aspx?NewsId=16374&CatId=122. [Đã truy cập 12 05 2024]. |
| 1. Đặng Thị Ánh Tuyết và Vũ Thái, “Ảnh hưởng và hậu quả của tin giả (Fake News) trên không gian mạng đến chính trị, kinh tế, xã hội,” 26 3 2023. [Trực tuyến]. Available: https://hdll.vn/vi/nghien-cuu---trao-doi/anh-huong-va-hau-qua-cua-tin-gia-fake-news-tren-khong-gian-mang-den-chinh-tri-kinh-te-xa-hoi.html. [Đã truy cập 12 5 2024]. |
| 1. Thùy An, “Fakebook tuyên chiến với Deepfake,” 9 1 2020. [Trực tuyến]. Available: https://vtv.vn/cong-nghe/facebook-tuyen-chien-voi-deepfake-20200109142546234.htm. |
| 1. Phú Lữ, “Dùng Deepfake để lừa đảo, tống tiền người nổi tiếng, doanh nhân,” 29 11 2023. [Trực tuyến]. Available: https://cand.com.vn/Cong-nghe/dung-deepfake-de-lua-dao-tong-tien-nguoi-noi-tieng-doanh-nhan-i715369/. |
| 1. Trọng Minh và Nguyễn Diệp, “Deepfake và những nguy cơ với an toàn thông tin Việt Nam,” 24 3 2024. [Trực tuyến]. Available: https://petrotimes.vn/   deepfake-va-nhung-nguy-co-voi-an-toan-thong-tin-tai-viet-nam-707998.html. |
| 1. Du Lam, “Bùng nổ nhu cầu dùng deepfake nói chuyện với người đã khuất,” 14 5 2024. [Trực tuyến]. Available: https://vietnamnet.vn/bung-no-nhu-cau-dung-deepfake-noi-chuyen-voi-nguoi-da-khuat-tai-trung-quoc-2280219.html. |
| 1. M. Westerlund, “The Emergence of Deepfake Technology: A Review,” 2019. [Trực tuyến]. Available: https://www.proquest.com/docview/   2329154005?sourcetype=Scholarly%20Journals. [Đã truy cập 15 5 2024]. |
| 1. NGỌC BÍCH, “Deepfake và lời cảnh tỉnh xã hội,” 2024. [Trực tuyến]. Available: https://cand.com.vn/Ho-so-mat/deepfake-va-loi-canh-tinh-xa-hoi-i727516/. [Đã truy cập 12 5 2024]. |
| 1. Cristian Vaccari và Andrew Chadwick, “Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News,” 2020. [Trực tuyến]. Available: https://   www.researchgate.net/publication/339385335\_Deepfakes\_and\_Disinformation\_Exploring\_the\_Impact\_of\_Synthetic\_Political\_Video\_on\_Deception\_Uncertainty\_and\_Trust\_in\_News. [Đã truy cập 13 5 2024]. |
| 1. Thùy Dung, “Chuyên gia RMIT kêu gọi “lật mặt” thách thức từ deepfake,” 2024. [Trực tuyến]. Available: https://www.rmit.edu.vn/vi/tin-tuc/tat-ca-tin-tuc/2024/feb/chuyen-gia-rmit-keu-goi-lat-mat-thach-thuc-tu-deepfake. [Đã truy cập 12 5 2024]. |
| 1. TỊNH ANH, “Deepfake: Ác mộng mới của các chiến dịch bầu cử,” 2024. [Trực tuyến]. Available: https://cuoituan.tuoitre.vn/deepfake-ac-mong-moi-cua-cac-chien-dich-bau-cu-2024032008550249.htm. [Đã truy cập 12 5 2024]. |
| 1. LÊ SAN, “Rủi ro từ công nghệ deepfake,” 2024. [Trực tuyến]. Available: https://nhandan.vn/rui-ro-tu-cong-nghe-deepfake-post797093.html. [Đã truy cập 12 5 2024]. |
| 1. Lê Anh Thực, Hoàng Hữu Trí, “DEEPFAKE: Khi thú vui công nghệ trở thành…mối đe dọa khó lường,” 2022. [Trực tuyến]. Available: https://hvctcand.bocongan.gov.vn/llct-xdll-cand/ly-luan-chinh-tri/deepfake-khi-thu-vui-cong-nghe-tro-thanhmoi-de-doa-kho-luong-4354. [Đã truy cập 12 5 2024]. |
| 1. Hằng Trần, ““Giải oan” cho Deepfake: Ưu điểm và nhược điểm trong triển khai digital marketing,” 26 6 2022. [Trực tuyến]. Available: https://advertisingvietnam.com/giai-oan-cho-deepfake-uu-diem-va-nhuoc-diem-trong-trien-khai-digital-marketing-p19958. [Đã truy cập 5 12 2024]. |
| 1. H. V. Đức Phan, “Làm sao tránh bị lừa đảo bằng AI, Deepfake trên không gian mạng?,” 18 4 2023. [Trực tuyến]. Available: https://vneconomy.vn/lam-sao-tranh-bi-lua-dao-bang-ai-deepfake-tren-khong-gian-mang.htm. [Đã truy cập 12 5 2024]. |