OUCH! Março de 2022

Aprenda uma nova habilidade de sobrevivência: detectar deepfakes

O que são Deepfakes?

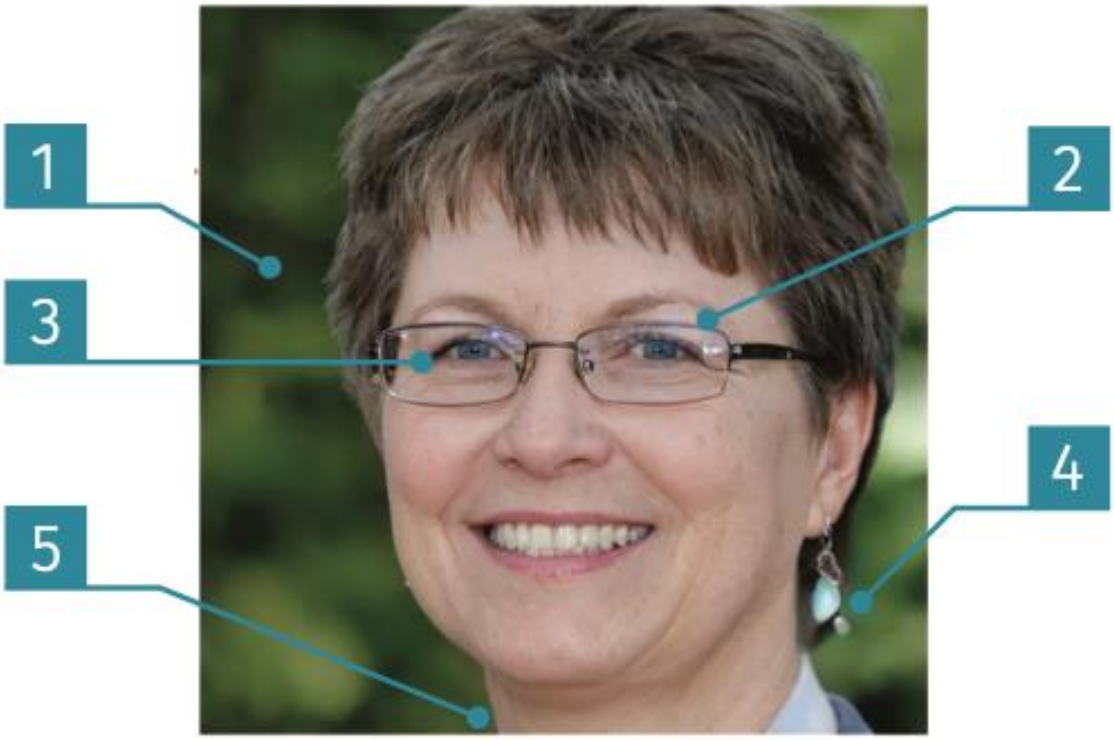
A palavra "deepfake" é uma combinação de "deep learning" e "fake". Deepfakes são montagens falsas imagens, vídeos ou gravações de áudio. Às vezes, as pessoas neles são geradas por computador, identidades falsas que parecem e soam como se fossem pessoas reais. Às vezes as pessoas são reais, mas suas imagens e vozes são manipuladas para fazer e dizer coisas que não fizeram ou nunca disseram. Por exemplo, um vídeo deepfake pode ser usado para fazer a montagem de uma celebridade ou político dizendo algo que nunca disse. Usando essas falsificações muito realistas, os atacantes cibernéticos podem criar uma realidade alternativa na qual você nem sempre pode confiar no que vê ou escuta.

Alguns deepfakes têm fins legítimos, como filmes que trazem atores falecidos de volta à vida para recriar um personagem famoso. Mas os atacantes cibernéticos estão começando a aproveitar o potencial dos deepfakes. Eles os empregam para enganar seus sentidos, para que possam roubar seu dinheiro, assediar pessoas, manipular eleitores ou opiniões políticas ou criar notícias falsas. Em alguns casos, eles até criaram empresas falsas compostas por funcionários deepfake. Você deve ser ainda mais cuidadoso com o que acredita ao ler notícias ou redes sociais diante desses ataques.

O FBI alerta que no futuro os deepfakes terão “um impacto maior e generalizado devido ao nível de sofisticação da mídia sintética utilizada”. Aprenda a identificar os sinais de um deepfake para se proteger dessas simulações extremamente realistas. Cada forma de deepfake– imagem estática, vídeo e áudio– tem seu próprio conjunto de falhas que podem denunciá-lo.

Imagens estáticas

O deepfake que você vê mais frequentemente é a foto falsa do perfil nas redes sociais. A imagem abaixo é um exemplo de deepfake do site thispersondoesnotexist.com. Abaixo da imagem temos cinco pistas diferentes de que isso pode ser um deepfake. Você vai perceber que essas pistas não são fáceis de detectar e podem ser difíceis de identificar:



1. Plano de fundo: o plano de fundo normalmente fica desfocado ou torto e pode ter uma iluminação irregular, como sombras exageradas apontando em várias direções.
2. Óculos: observe atentamente a conexão entre a armação e as hastes perto da têmpora. Os deepfakes normalmente têm conexões incompatíveis com tamanhos ou formas ligeiramente diferentes.
3. Olhos: as fotos com deepfake atualmente usadas para fotos de perfil falsas parecem ter os olhos no mesmo ponto do quadro, resultando no que alguns chamam de "olhar deepfake".
4. Joias: os brincos podem estar disforme ou presos de um jeito estranho. Os colares podem estar encorporados na pele.
5. Colares e ombros: Os ombros podem estar deformados ou ser incompatíveis. Os colares podem ser diferentes em cada lado.

Vídeo

Pesquisadores do Instituto de Tecnologia de Massachusetts, MIT, desenvolveram uma lista de perguntas para ajudar você a descobrir se um vídeo é real, observando que os deepfakes normalmente não conseguem "representar totalmente a física natural" de uma cena ou iluminação.

1. Bochechas e testa: A pele tem uma aparência muito lisa ou enrugada? A idade da pele é semelhante à idade do cabelo e dos olhos?
2. Olhos e sobrancelhas: As sombras aparecem nos lugares que você esperaria?
3. Óculos: Tem algum brilho? Muito brilho? O ângulo do brilho muda quando a pessoa se move?
4. Pelos faciais: Os pelos faciais parecem reais? Os deepfakes podem adicionar ou remover bigode, costeletas ou barba.
5. Pintas no rosto: As pintas parecem ser de verdade?
6. Pestanejar: A pessoa pisca o suficiente ou demais?
7. Tamanho e cor dos lábios: O tamanho e a cor combinam com o resto do rosto da pessoa?

Áudio/Voz

Pesquisadores afirmam que tecnologias como espectrogramas podem mostrar quando as gravações de voz são falsas. Mas, a maioria de nós não tem o luxo de um analisador de voz quando um invasor nos lista. Preste atenção a uma voz, tom ou emoção voz invariável e estranha, além da falta de ruído de fundo. As falsificações de voz podem ser difíceis de detectar. Se você receber uma chamada estranha de uma organização legítima, poderá verificar se a chamada é verdadeira desligando primeiro e depois ligando de volta para essa organização. Certifique-se de usar um número de telefone confiável, como um número de telefone que você já tenha em sua lista de contatos, um número de telefone impresso em uma conta ou extrato da organização ou o número de telefone no site oficial da organização.

Conclusão

Esteja ciente de que os invasores estão usando os deepfakes ativamente. Eles podem criar contas falsas nas redes sociais para se conectar ou criar vídeos falsos para manipular a opinião pública. Alguns estão até vendendo seus serviços na dark web para que outros invasores possam utilizá-los. Não esperamos que você se torne um especialista em deepfake, mas se aprender o básico para identificar as montagens, vai conseguir se defender muito bem. Se você suspeitar que detectou um deepfake, denuncie ao site ou fonte que hospeda esse conteúdo.

Editor convidado

Kerry Tomlinson (@KerryTNews) é um jornalista que cobre notícias cibernéticas da Ampere News e um profissional certificado de conscientização de segurança do SANS. Sua missão é traduzir o que está acontecendo no mundo digital para pessoas de todos os níveis de conhecimento com notícias convincentes e perspicazes, além de apresentações chamativas.

Recursos

Engenharia Social: https://www.sans.org/newsletters/ouch/social-engineering-attacks/
Você consegue detectar a montagem? (Ampere News): https://www.amperesec.com/news/can-you-spot-the-fake
Teste de detecção de deepfake do MIT (MIT): https://detectfakes.media.mit.edu/
Identifique o deepfake: https://www.spotdeepfakes.org/en-US

Traduzido para a Comunidade por: David Boldrin