Team Name:
Bugs_Slayer

Project Name:
AuthentiFake

Team Members:
Jireh Baillo
Arthur Chen
Wesley Ip
Jordan Karaze
Marlon Pimentel
Hitesh Verma
Sanjula Yella
IBM Office Location: Austin, TX
Team Number: 38

GitHub: https://github.com/wip34/BugSlayer_AuthentiFake

AuthentiFake Summary

In a world where a single fake image shared on social media can cause rifts between entire ethnic groups, we need a means to authenticate images quickly and efficiently. In this political and economic climate, news has been circulating faster than ever; unfortunately, attention spans have been getting proportionally shorter. Convenience now rules every aspect of our lives, blurring the lines between fact and fiction. This is the era of fake news, and with the power to influence hundreds of people at our fingertips, it has never been easier to circulate fake, and potentially dangerous images. With Deepfake technology becoming increasingly advanced, fake nude photos tainting reputations are the least of our worries. Malicious users can (and have) doctor images to show presidential candidates in compromising situations, a religion or culture being disrespected by an opposing group, and healthy ice caps where there are none. This technology has the potential to be so catastrophic that the Pentagon has taken a special interest in finding a remedy.

Our product, AuthentiFake, brings you a first-of-its-kind, result-oriented solution to combat fire with fire (AI with AI). Using IBM Watson Studio, Blockchain Platform, IKS Cluster, Machine Learning our product will warn users trying to upload an image on partnered social media sites that an image has been altered. Using the open source blockchain framework, Hyperledger, and Artificial Intelligence capabilities similar to Deepfake, AuthentiFake will scan an image for mismatched pixel anomalies and trigger an event. This causes relevant information including hash value, edits, social media platform, and geographical location to be stored on a permissioned blockchain that only our clients (media companies) and IBM can access. We believe that collaboration, rather than competition, is the key to a brighter future; the effectiveness of AuthentiFake is dependent on the number of popular social media sites that implement it. The final output is a minimally intrusive message that notifies the user that the image they are trying to share with the world is not what it seems.

AuthentiFake will lead the market in the fight against fake images. The trajectory of Deepfake technology and usage reveals that it will reach problematic heights in roughly one year as progress continues. Within this time, with widespread implementation of our microservice, AuthentiFake can be scaled up to include video detection and more advanced hashing algorithm technology to increase transparency and accountability.