**Detecting Fake Images on social media using Machine Learning**

**Abstract:**

In this technological era, social media has a major role in people’s daily life. Most people share text, images, and videos on social media frequently (e.g. Twitter, Snapchat, Facebook, and Instagram). Images are one of the most common types of media share among users on social media. So, there is a need for monitoring of images contained in social media. It has become easy for individuals and small groups to fabricate these images and disseminate them widely in a very short time, which threatens the credibility of the news and public confidence in the means of social communication. This research attempted to propose an approach to extracting image content, classify it and verify the authenticity of digital images and uncover manipulation. Instagram is one of the most important websites and mobile image sharing applications on social media. This allows users to take photos, add digital photographic filters and upload pictures. There are many unwanted contents in Instagram's posts such as threats and forged images, which may cause problems to society and national security. This research aims to build a model that can be used to classify Instagram content (images) to detect any threats and forged images. The model was built using deep algorithms learning which is Convolutional Neural Network (CNN), Alexnet network and transfer learning using Alexnet. The results showed that the proposed Alexnet network offers more accurate detection of fake images compared to the other techniques with 97%. The results of this research will be helpful in monitoring and tracking in the shared images in social media for unusual content and forged images detection and to protect social media from electronic attacks and threats.