

FORENSISCHE BILDANALYSE

28.08.2024

Analyst: Anatol Maier M.Sc.

Software: Neuraforge Deepfake-Detection



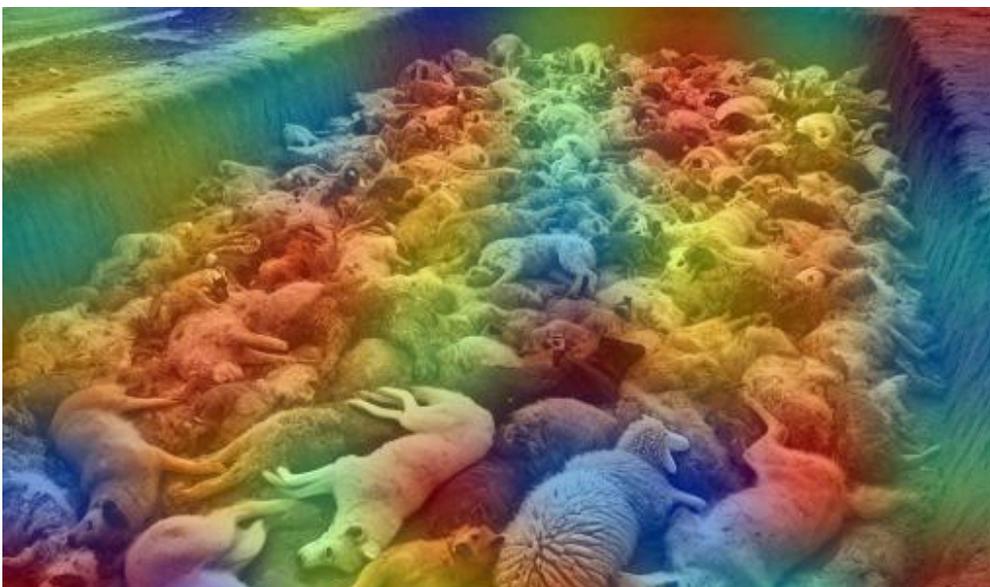
Bildquelle:

https://x.com/vegan_earth/status/1821814389827100938

Ergebnis und Interpretation:

Die Analyse des Bildes ergibt **mit sehr hoher Wahrscheinlichkeit**, dass **dieses Bild KI-generiert** ist. Das Modell, mit dem dieses Bild generiert wurde, ist durch die **starke Kompression**, die mit Verbreitung über X einher geht, nicht eindeutig zuzuordnen. Jedoch überlagern sich die Spuren, die für KI-generierten Inhalt sprechen, nicht mit den Kompressionsartefakten, weswegen hier auch eine hohe Modellzuverlässigkeit interpretiert werden kann.

Die unten stehende Heatmap zeigt die relevantesten Spuren für das Modell, welche auf KI-content hindeuten. Unser Ansatz sucht nach statistischen Spuren unter anderem im Frequenzbereich und nicht nach semantischen Fehlern im Bild. Daher können sich zwar die für unser Modell relevanten Spuren mit visuell erkennbaren „Fehlern“, wie fehlende (oder zu viele) Pfoten oder fehlendes Auge, überlagern, müssen jedoch nicht.



Heatmap: Von Blau zu rot werden Bereiche im Bild markiert, die laut Analyse Artefakte von KI Bildgeneratoren enthalten. Blaue Bereiche sind für die Einordnung als Fake weniger relevant - rote und gelbe Bereiche sind sehr relevant.

Zusätzlich lassen sich auch noch durch **visuelle Inspektion typische „Fehler“**, die KI-Generatoren aktuell noch machen, erkennen.

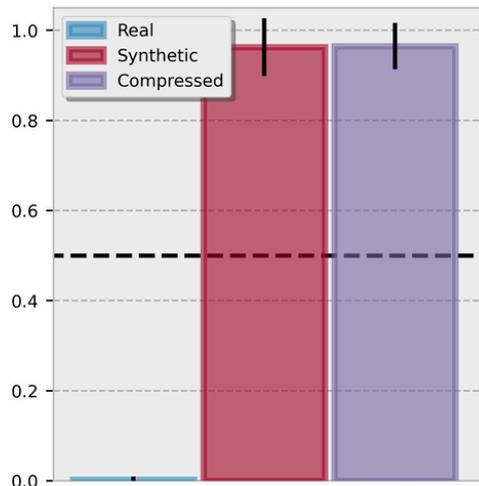
Optische Fehler in KI-Generierten Bildern entstehen meist weil die KI danach strebt, runde oder glatte Übergänge im Bild herzustellen. Da all die Informationen aber nur implizit aus Millionen von Trainingsbildern erlernt sind und kein explizites Verständnis von Körperformen, Proportionen und Bildaufbau besteht, entstehen diese Art von Fehlern im Bild.



Beim untersuchten Bild fällt auf, das einige Hunde Zyklopen zu sein scheinen oder komplette Gesichter fehlen. Auch stimmen manche Proportionen nicht. Vor allem im hinteren Bereich des “Massengrabes” verschmelzen die Körper regelrecht in einander, was oft auf KI-Fake hindeutet.

Auch sind hier einige **Unregelmäßigkeiten in der generellen Bildsemantik** zu erkennen. Zum Beispiel sind die aufgeworfenen Erdhügel im Hintergrund viel zu klein für die Größe des hier dargestellten Loches.

Alles in allem kann man mit an Sicherheit grenzender Wahrscheinlichkeit sagen, dass das untersuchte **Bild vollständig KI generiert** ist.



Grafik der Fake-Analyse:
Das Modell zeigt beim untersuchten Bild eine hohe Konfidenz in den Klassen “Synthetic” und “compressed” an.

Für Rückfragen stehen wir jederzeit gerne zur Verfügung

Anatol Maier und Anika Gruner Neuraforge GbR



info@neuraforge.de
www.neuraforge.de