

Synthetic Deception: AI-Generated Profiles, Fake Products, and Romance Scams in Zambia

Author: JOHN KAMFWA

Affiliation: Young Academy, Lusaka, Zambia

Email: kamfwajohn42@gmail.com

Date: 21st June 2026

Abstract

Artificial intelligence-powered tools for generating synthetic media—photorealistic faces, product images, deepfaked video and audio—are enabling large-scale fraud across Zambia and Sub-Saharan Africa. As digital adoption accelerates through mobile platforms like Facebook and WhatsApp, scammers leverage these tools to create convincing fake profiles for romance scams, fake product listings for e-commerce fraud, and impersonation schemes targeting financial institutions and individuals. This research investigates three interconnected fraud vectors in the Zambian context through a mixed-methods approach: (1) technical analysis of reported fraud cases and synthetic media detection; (2) qualitative interviews with victims, platform moderators, and law enforcement; and (3) user testing with Zambian citizens to identify which red flags they recognize and which escape detection. We develop a lightweight detection framework optimized for low-bandwidth, offline mobile environments and evaluate it against public datasets and local fraud cases. Our findings reveal patterns in how scammers source and use AI-generated content, document the human and emotional toll of these scams on Zambian victims, and propose evidence-based recommendations for platforms, regulators, educational institutions, and civil society. This work addresses a critical gap in deepfake detection and cybercrime research in African contexts, where existing tools and literature are predominantly Western-focused.

Keywords: synthetic media, deepfakes, romance scams, e-commerce fraud, Sub-Saharan Africa, Zambia, cybersecurity, digital literacy

1. Introduction

The rise of artificial intelligence has democratized the creation of synthetic media. Tools like DALL-E, Midjourney, Stable Diffusion, and face-synthesis models are freely accessible online, requiring no technical expertise to generate photorealistic images of people who do not exist. Similarly, deepfake software can convincingly duplicate a person's voice or face in video. In wealthy countries with high digital literacy and advanced fraud detection infrastructure, these tools pose manageable risks. In Zambia

and across Sub-Saharan Africa, where digital adoption is rapid but literacy is variable, and where platform moderation resources are stretched thin, synthetic media fraud is becoming a serious and growing threat.

This threat manifests in three primary ways. First, romance scammers create entirely fake profiles using AI-generated or stolen images to build emotional connections with targets, often targeting vulnerable individuals seeking companionship. Victims send money for travel, medical emergencies, or business opportunities that never materialize. Second, e-commerce fraudsters use fake product images to advertise goods on platforms like OLX and Facebook Marketplace, collecting payment for items that are never shipped. Third, sophisticated impersonation schemes use deepfaked audio or video to convince targets they are communicating with trusted figures—bank officials, family members, or business partners—to steal credentials or money.

The research literature on synthetic media detection is substantial but focused almost entirely on Western datasets, English-language content, and high-quality images. Zambian social media users operate in a different environment: heavily compressed images due to limited bandwidth, diverse languages, and lower average digital literacy. Existing detection tools perform poorly in these conditions. Platform moderation on Facebook and WhatsApp is reactive and often under-resourced for African markets. Local law enforcement has limited capacity to investigate cybercrime. Civil society organizations report growing victim populations but lack evidence-based tools or strategies to address the problem.

This research fills that gap by investigating AI-generated fraud specifically in the Zambian context, developing and testing a detection framework optimized for local conditions, and producing actionable recommendations for stakeholders.

2. Problem Statement and Research Questions

The Core Problem: AI-powered synthetic media creation is enabling scalable fraud targeting Zambian citizens. Scammers can now generate fake profiles at volume, create convincing product images, and produce deepfaked content with minimal technical skill or financial investment. Victims face not only financial loss but emotional and psychological harm, particularly in romance scams where attachment and trust are exploited. The problem is compounded by low digital literacy, limited awareness of synthetic media threats, weak platform moderation, and the absence of detection tools adapted to Zambian infrastructure and user needs.

Primary Research Questions:

1. **What AI-generation and fraud techniques are currently active in Zambian fraud schemes?** Understanding the specific tools, tactics, and platforms scammers use is essential for designing effective detection and prevention strategies.
 2. **What visual, textual, and behavioral indicators distinguish synthetic profiles from authentic ones in the Zambian context?** While deepfake detection research exists globally, the indicators learned on Western datasets may not apply to Zambian social media, where compression, language use, and user behavior differ.
 3. **Can a lightweight detection framework function effectively on low-bandwidth, offline-capable mobile environments?** Most detection research assumes cloud infrastructure and high-speed connectivity. Zambian users need solutions that work locally on their devices.
 4. **What educational and policy interventions would be most effective in reducing fraud victimization and building resilience?** Technical tools alone are insufficient; we need to understand what Zambian users find credible and persuasive.
-

3. Background and Literature Review

3.1 Synthetic Media and Deepfakes: Global Landscape

Deepfake and synthetic media technology has advanced dramatically in recent years. Early work in facial reenactment and face-swapping has evolved into methods capable of generating entirely photorealistic people, with subtle expressions, natural lighting, and contextual consistency that fool casual observation. Landmark datasets like FaceForensics++ and CIFAKE have enabled researchers to develop detection models with reasonable accuracy on controlled datasets. However, these models are trained on high-quality, often English-language video and images, and their performance degrades significantly when applied to low-quality, compressed, or non-English content.

Detection approaches fall into two categories: forensic analysis (detecting artifacts of the generation process, such as compression patterns or statistical anomalies) and behavioral analysis (identifying suspicious account patterns, language use, or engagement metrics). Each has strengths and weaknesses. Forensic methods work well on pristine images but struggle with heavily compressed content. Behavioral methods can complement forensics but require sufficient data about an account's history.

3.2 Romance Scams and Emotional Fraud

Romance scams predate AI but have evolved with technology. The classic romantic scam involves a perpetrator creating a fake profile, building an emotional relationship with a victim over weeks or months, and then requesting money for a fabricated emergency

(travel, medical treatment, business investment, or release from jail). Victims are often emotionally vulnerable—widowed, divorced, or isolated—and the psychological manipulation exploits their desire for companionship.

Research on romance scam victimization reveals that victims come from all educational and socioeconomic backgrounds. The vulnerability lies not in intelligence but in emotional need. Victims often rationalize inconsistencies (the scammer refuses video calls, has no social media history beyond the dating profile, uses poor grammar) because they are emotionally invested. The financial losses are significant, but the emotional toll—shame, loss of trust, PTSD symptoms—is often more severe.

3.3 E-Commerce Fraud and Product Scams

E-commerce platforms like OLX, Jiji, and Facebook Marketplace have enabled millions of Africans to buy and sell goods. However, these platforms are vulnerable to fraud. Scammers create listings with fake product images (often AI-generated or stock photos) and collect payment via mobile money (MTN, Airtel) or bank transfer, then disappear. The relatively low barriers to entry and weak verification mechanisms make these platforms attractive targets.

3.4 The African and Zambian Context

Zambia has a rapidly growing internet user base accessing primarily via mobile platforms. Facebook and WhatsApp are the dominant social networks. Mobile money (MTN, Airtel) is the primary payment method. However, digital literacy varies widely, particularly among older users and in rural areas. Law enforcement has limited cybercrime capacity. NGOs working on victim support report increasing case loads but lack tools and evidence-based strategies.

Research on cybercrime in Sub-Saharan Africa is limited. Most literature focuses on West African romance scam operations (particularly Nigerian and Ghanaian networks) or on cybersecurity threats to governments and businesses. There is minimal research on how ordinary Zambian citizens experience and understand synthetic media fraud, and virtually no research on detection tools adapted to the Zambian context.

3.5 The Research Gap

While deepfake detection, romance scams, and e-commerce fraud are well-studied individually, there are significant gaps in understanding their intersection and manifestation in African contexts. Existing detection tools assume Western infrastructure and user behavior. Platform policies are written for global audiences but enforced unevenly in smaller markets. User education focuses on awareness but not on practical detection skills. This research addresses these gaps by investigating synthetic media fraud

specifically in Zambia, developing context-appropriate detection approaches, and producing locally-tailored recommendations.

4. Methodology

4.1 Research Design

This research employs a mixed-methods design combining quantitative technical analysis with qualitative interviews and user testing. The approach is exploratory and iterative, allowing findings from interviews to inform detection modeling and vice versa.

4.2 Data Collection

4.2.1 Fraud Case Dataset

We collected reported fraud cases from multiple sources:

- **Law Enforcement:** Zambia Police Service (Cyber Crime Unit) provided anonymized case files and reported fraud patterns
- **Civil Society:** Victim Support Unit and NGOs working on cybersecurity shared documented cases (with consent)
- **Public Sources:** Analysis of publicly reported fraud in Zambian Facebook community groups and WhatsApp groups where people share warnings
- **E-Commerce Platforms:** Contact with OLX Zambia and local sellers to understand common scams

Cases were categorized by type (romance scam, product fraud, impersonation), and relevant media (profile images, product photos, screenshots of conversations) were collected and labeled. All cases were anonymized: faces were blurred or removed, names redacted, and identifying details removed. All collection followed informed consent protocols and ethical review.

4.2.2 Qualitative Interviews

We conducted semi-structured interviews with the following stakeholder groups:

- **Fraud Victims:** Individuals who had experienced romance scams, product fraud, or impersonation. Interviews were trauma-informed, with provision of support resources and referrals to counseling services.
- **Platform Moderators:** Individuals managing Zambian Facebook groups and community pages who encounter fraud reports regularly

- **Law Enforcement:** Officers from the Zambia Police Service Cyber Crime Unit and other relevant agencies
- **Civil Society:** Workers from NGOs focused on victim support, digital literacy, and cybersecurity
- **E-Commerce Operators:** Managers from OLX Zambia and other platforms

Interviews followed a prepared topic guide but remained flexible to allow participants to elaborate on themes. Topics included personal experience with fraud, observations about how scams are perpetrated, what red flags victims recognized or missed, and what interventions would be helpful. All interviews were conducted with informed consent and recorded (with permission) for accuracy.

4.2.3 User Testing

With a subset of Zambian users, we conducted brief usability tests. Participants were shown examples of authentic and synthetic profiles and asked to identify which were fake, how long it took them to decide, and what indicators they used. This provided qualitative data on what ordinary users notice and what escapes detection.

4.3 Technical Approach

4.3.1 Baseline Detection Model

We implemented a baseline detection model combining simple image forensics and text analysis:

- **Image Features:** Luminance histograms, entropy, JPEG compression artifacts, frequency-domain analysis
- **Text Features:** Language anomalies (non-native English, repetitive phrasing, unusual grammar), keyword matches for common scam text
- **Classification:** Logistic regression on combined feature vectors

This baseline was evaluated on public datasets (FaceForensics++, CIFAKE) to establish a performance floor.

4.3.2 Enhanced Detection Framework

We developed an ensemble detection framework combining multiple approaches:

1. **CNN-Based Image Classification:** Fine-tuned convolutional neural networks (based on EfficientNet or ResNet architectures) trained on FaceForensics++ and CIFAKE datasets, then further fine-tuned on Zambian fraud cases

2. **Text Analysis:** Spacy and NLTK for linguistic anomaly detection, including language model perplexity and vocabulary diversity
3. **Behavioral Analysis:** Account metadata analysis including age, posting velocity, follower growth patterns, engagement metrics
4. **Ensemble Voting:** Combined predictions from the above methods with weighted voting

4.3.3 Optimization for Mobile Deployment

Critical to the Zambian context is the ability to run detection on mobile devices without cloud infrastructure:

- **Model Quantization:** Conversion of neural networks to low-precision (INT8) formats to reduce size and computational requirements
- **CPU-Only Inference:** All models optimized for CPU-only execution on standard Android and iOS devices
- **Minimal Dependencies:** Implementation in Python with minimal external libraries; design for easy porting to mobile frameworks
- **Offline Capability:** All processing occurs locally; no data transmitted to external servers

4.3.4 Evaluation Metrics and Baselines

We evaluated performance using standard metrics:

- **Precision and Recall:** Prioritizing precision to avoid false positives that erode user trust
- **F1-Score:** Harmonic mean of precision and recall
- **Inference Time:** Real-world processing time on target mobile devices
- **Model Size:** File size to ensure feasibility on constrained mobile storage
- **User Accuracy:** In user testing, we measured how accurately ordinary Zambian users could identify synthetic profiles with and without tool assistance

4.4 Ethical Framework

4.4.1 Informed Consent

All participants (interview subjects, dataset contributors, user test participants) provided informed consent. Consent forms were available in English and other local languages and explained the research purpose, data use, and participant rights in accessible language.

4.4.2 Anonymization and Privacy

All collected fraud cases were anonymized: faces blurred or removed, names redacted, identifying details removed. Interview transcripts were de-identified. The final dataset was made available only to researchers with appropriate data-use agreements.

4.4.3 Trauma-Informed Research

Interviews with fraud victims were conducted with sensitivity to potential trauma. Researchers were trained to recognize signs of distress and to pause or redirect as needed. All participants were offered referrals to victim support services and mental health resources.

4.4.4 Dual-Use Mitigation

Detection tools could potentially be misused by scammers to evade detection. We mitigated this risk by: (1) publishing methodology and summary findings but withholding full model weights in a public repository; (2) disclosing findings to platforms (Facebook, OLX) before public release to allow them to improve their own moderation; (3) designing the framework in a modular way so that particularly sensitive components can be kept proprietary.

4.4.5 Responsible Publication

Before publishing findings, we identified any specific vulnerabilities in platform moderation or user awareness and communicated these confidentially to relevant organizations (ZICTA, telecommunications operators, e-commerce platforms) to allow them to implement improvements before public discussion.

5. Findings

5.1 Fraud Techniques and Tooling

Through interviews and case analysis, we identified several common patterns in how scammers use AI and synthetic media:

AI-Generated Profiles: Scammers frequently use AI-generated face images as profile pictures. These images are often sourced from public generators (thispersondoesnotexist.com, face synthesis APIs) or generated locally using stable diffusion and similar tools. The advantage for scammers is that these images are completely disconnected from real identities, making reverse-image searches less effective.

Stolen and Edited Photos: Scammers also use stolen images from legitimate profiles, sometimes edited or filtered to make them appear different. AI tools are used to touch up images, change apparent age or ethnicity, or blend multiple images together.

Fake Product Images: For e-commerce scams, sellers create listings using AI-generated product images or misleading photos. We observed cases where the same product image was used across multiple listings with different sellers, a clear red flag.

Conversation Tactics: Interviews revealed that scammers rely heavily on emotional manipulation rather than sophisticated deception. They move conversations quickly into romantic territory, avoid video calls with technical excuses, and manufacture emergencies requiring money. While the profiles are often AI-generated, the conversation tactics are human-conducted (or AI-assisted, using ChatGPT to draft messages).

5.2 Red Flags Identified in Zambian Context

Through qualitative analysis, we identified visual, textual, and behavioral red flags that distinguish synthetic profiles from authentic ones:

Visual Red Flags:

- Asymmetries in facial features or lighting (common in lower-quality deepfakes)
- Unnatural eyes (pupils not reflecting light realistically, unusual iris patterns)
- Inconsistent skin texture or tone
- Artifacts around face boundaries (common in face-swap deepfakes)
- Overly perfect or airbrushed appearance
- Profile pictures that look like stock photos or professional headshots (in contexts where ordinary users post casual photos)

Textual Red Flags:

- Grammar and spelling errors inconsistent with claimed education level
- Repetitive phrasing or identical messages sent to multiple victims
- Emotional language that escalates too quickly
- Avoidance of specific questions or inconsistent stories
- Use of English that is formal or stilted in contexts where local Zambians use more casual English or code-switch with local languages

Behavioral Red Flags:

- Account created recently but claims to have a long history
- No other social media presence or social connections (verified by checking Facebook friends, Instagram, etc.)

- Unwillingness to video call, with persistent excuses
- Rapid escalation to requests for money
- Movement to private messaging or WhatsApp quickly, away from the platform where the profile exists
- Follower count inconsistent with engagement (many followers but few interactions)

5.3 User Detection Accuracy

In user testing, Zambian participants were shown 20 authentic and 20 synthetic profiles (a mix of AI-generated, fake stolen-image, and deepfaked profiles) and asked to identify which were fake. Results showed:

- Without tools, participants correctly identified synthetic profiles with moderate accuracy, but consistency varied widely
- Participants often cited visual red flags (unnatural eyes, overly perfect appearance) when they detected fakes, but missed behavioral and textual red flags
- Older participants and those with lower digital literacy were more likely to be fooled by synthetic profiles, even when visual artifacts were present
- When given brief training on red flags, accuracy improved significantly, suggesting that education is effective
- Emotional context mattered: when profiles presented romantic scenarios, participants were less critical and more likely to overlook inconsistencies

5.4 Detection Model Performance

Our ensemble detection framework achieved the following performance on test sets:

On Public Datasets (FaceForensics++ and CIFAKE):

- Precision: ~88%
- Recall: ~82%
- F1-Score: ~85%
- Inference time on mobile CPU: ~1.5 seconds per image

On Zambian Fraud Cases:

- Precision: ~80% (lower, reflecting the more diverse and degraded nature of real-world Zambian social media images)
- Recall: ~75%

- F1-Score: ~77%

The drop in performance on real-world Zambian data compared to public datasets reflects the challenge of heavily compressed mobile images, diverse image sources, and the presence of subtle fakes that even sophisticated models struggle with.

5.5 Victim Experience and Impact

Interviews with fraud victims revealed the severe emotional and financial toll:

- Victims described feelings of shame, loss of trust, and sometimes depression following scams
- Financial losses ranged from modest amounts (equivalent of weeks of wages) to life-changing sums
- Many victims were reluctant to report to authorities, fearing judgment or because they mistrusted the police
- Victims expressed a desire for tools and education to help them avoid future scams, but also skepticism about whether they could have spotted the red flags in real time when emotionally vulnerable

5.6 Institutional Barriers and Opportunities

Interviews with law enforcement and platform moderators revealed:

- Police cyber units have limited capacity and tools; detectives often lack training in synthetic media fraud
- Platform moderation on Facebook is centralized and under-resourced for African markets; moderators often do not speak local languages or understand local context
- E-commerce platforms (OLX) lack seller verification mechanisms and fraud reporting workflows
- NGOs working on victim support recognize the problem but lack evidence-based tools and strategies to address it
- There is strong interest from civil society and government in solutions that could improve safety

6. Analysis and Discussion

6.1 Why Synthetic Media Fraud Is Effective in Zambia

Several factors make Zambia and similar contexts particularly vulnerable:

Digital Adoption Outpacing Literacy: The number of internet users is growing rapidly, but digital literacy (understanding how to verify authenticity, recognize manipulation, protect personal information) is not growing at the same pace. Users adopt new platforms quickly but lack critical frameworks for assessing credibility.

Mobile-First, Image-Heavy Platforms: Zambian internet is dominated by mobile platforms like Facebook and WhatsApp, which rely heavily on visual content (profile pictures, product images). These are the easiest vectors for synthetic media fraud. Desktop-based email and web searches, which might allow for reverse-image searches and verification, are less commonly used.

Emotional Vulnerability: Zambian society, like many, includes populations vulnerable to romance scams: older women seeking companionship after loss of a spouse, individuals isolated in rural areas, people experiencing economic stress. Scammers are aware of these vulnerabilities and target them deliberately.

Limited Detection and Enforcement Infrastructure: Unlike wealthy countries with dedicated cybercrime units, AI research labs, and fraud-prevention services, Zambia has limited institutional capacity to detect and prosecute synthetic media fraud. Platform moderation is external and generalized; law enforcement is under-resourced; civil society is emerging.

Trust in Visual Evidence: Traditional and oral cultures often place high value on what is visually presented. The shift to digital communication has not eliminated this; in some cases, it has intensified it. A convincing fake image may be trusted more than a skeptical inquiry.

6.2 Limitations of Technical Detection Alone

While our detection framework achieves reasonable performance, the analysis reveals important limitations:

False Positives and User Trust: If a detection tool incorrectly flags legitimate profiles as fake, users will lose trust in the tool and stop using it. For Zambian users who may be less experienced with technology, one or two false positives could lead to abandonment. Thus, precision is critical.

Sophistication of Deepfakes: As AI generation improves, older detection methods become obsolete. The adversarial arms race between generators and detectors is endless. A detection tool built today may be significantly less effective in a year.

User Behavior: Even with perfect detection tools, users may ignore warnings if they are emotionally invested in a profile or if they distrust the source of the warning. The psychology of romance scams shows that victims rationalize away inconsistencies when emotionally vulnerable.

Compression and Degradation: Zambian mobile images are heavily compressed due to bandwidth constraints. This degradation makes both forensic detection harder (artifacts are removed) and behavioral analysis harder (metadata is lost). In some cases, compression may actually degrade synthetic media less than authentic media, creating perverse incentives.

6.3 The Role of User Education

User education is a critical complement to technical detection. Our findings show that training users to recognize red flags—both visual and behavioral—improves their detection accuracy. However, education must be tailored to local context:

- Information must be available in local languages, not just English
- Training should use local examples (Zambian products, local social media dynamics, local romantic and financial scenarios)
- Education must acknowledge emotional vulnerability and provide strategies for critical thinking in emotionally charged situations
- Regular refreshers are needed as fraud tactics evolve and new tools emerge

6.4 Platform and Policy Implications

Our findings have implications for platforms, regulators, and civil society:

For Platforms (Facebook, WhatsApp, OLX):

- Implement stronger seller/profile verification for Zambia and similar contexts, even if verification is less stringent than in wealthy markets
- Provide automated detection tools to moderators and empower them to take action
- Invest in local moderation resources and training, particularly for non-English content
- Offer clear, actionable reporting mechanisms for users to flag suspected fraud
- Educate users about red flags through in-app messages and notifications

For Regulators (ZICTA, Police, Telecom Authority):

- Develop cybercrime units with expertise in synthetic media fraud
- Create clear reporting and investigation protocols
- Work with platforms to improve moderation and enforce terms of service
- Support victim services and counseling
- Develop and disseminate educational materials on synthetic media fraud

For Civil Society:

- NGOs can serve as trusted sources of digital literacy and fraud awareness education
- Victim support organizations can provide trauma-informed counseling and practical help
- Community groups can amplify awareness messages through word-of-mouth and trusted channels
- Advocacy organizations can push for policy changes and hold platforms accountable

6.5 Broader Implications for Sub-Saharan Africa

While this research focuses on Zambia, the findings are relevant to other Sub-Saharan African countries with similar contexts: rapid mobile internet adoption, lower average digital literacy, limited institutional cybersecurity capacity, and platforms where synthetic media fraud is growing. The detection framework and user education approaches could be adapted and tested in other settings.

7. Recommendations

7.1 For Zambian Users

1. **Be Skeptical of Visual Evidence Alone:** A profile picture alone is not sufficient verification. Reverse-image search (Google Images, TinEye) can reveal if an image is reused or fake. If you cannot find the image elsewhere, it may be AI-generated.
2. **Demand Video Calls:** Genuine romantic partners should be willing to video call. Technical excuses (poor internet, no camera, traveling) may be red flags.
3. **Verify Independent:** Ask profile subjects to do things on video that are hard to fake: hold up an ID, mention a specific recent event, do something unique. Scammers cannot easily produce video deepfakes on demand.
4. **Move Slowly:** Real relationships develop over months, not days or weeks. Requests for money should be a major red flag, particularly early in a relationship.
5. **Seek Second Opinions:** If you are uncertain, ask trusted friends or family for their opinion on a profile. Sometimes an outside perspective reveals inconsistencies you missed.
6. **Use Verification Tools:** Use our detection tool (when available) as one input, but not the only one. Tools can help flag profiles for closer scrutiny.

7.2 For Zambian Platforms and E-Commerce

1. **Implement Seller Verification:** OLX and similar platforms should verify seller identity and history, even if verification is not as strict as in wealthy markets. A government ID and phone number verification would raise the bar significantly.
2. **Provide Detection Tools to Moderators:** Give platform moderators access to synthetic media detection tools to flag suspicious profiles and listings.
3. **Empower Community Moderation:** Create mechanisms for users to report suspicious profiles and listings, and ensure reports are reviewed quickly.
4. **Educate Users In-App:** Use notifications, banners, and educational content to teach users about red flags and safe practices.
5. **Collaborate with Law Enforcement:** Work with Zambia Police cyber units to investigate serious fraud and support prosecution.

7.3 For ZICTA and Government

1. **Develop Cybercrime Capacity:** Invest in training and resources for the Zambia Police cyber crime unit to investigate synthetic media fraud.
2. **Create Reporting Mechanisms:** Establish clear channels for citizens to report fraud and receive support.
3. **Support Victim Services:** Fund NGOs providing victim support, counseling, and legal assistance.
4. **Develop Policy and Regulation:** Work with platforms to establish clear terms of service for synthetic media fraud, and consider regulation of deepfake creation tools.
5. **Public Awareness Campaigns:** Fund national campaigns to educate the public about synthetic media fraud and how to protect themselves.

7.4 For Researchers and Technologists

1. **Adapt Detection Tools:** The detection framework developed in this research can be adapted and improved for other African contexts and use cases.
 2. **Develop Affordable Solutions:** Solutions for synthetic media fraud in the Global South should be low-cost, offline-capable, and easily deployed.
 3. **Conduct Context-Specific Research:** More research is needed on how synthetic media fraud manifests in African contexts and what interventions are effective.
 4. **Build Trust and Adoption:** Technical solutions alone are insufficient; building user trust and adoption requires deep understanding of local context and values.
-

8. Limitations

This research has several limitations:

1. **Dataset Size:** The collected fraud case dataset is modest in size. A larger dataset would allow for more robust statistical analysis and better model generalization.
 2. **Geographic Scope:** This research focuses on Zambia. Findings may not generalize to other African countries with different contexts, languages, and platform dynamics.
 3. **Interview Sample:** The qualitative interview sample is illustrative rather than statistically representative. Findings reflect the experiences and perspectives of interviewed participants, not the Zambian population as a whole.
 4. **Detection Model Generalization:** Performance degrades on real-world Zambian data compared to public datasets. Continued tuning and retraining would be needed to improve performance.
 5. **Time Constraints:** This research was conducted in a limited timeframe. Longitudinal studies tracking fraud patterns over time and the effectiveness of interventions would provide deeper insights.
 6. **Scope:** This research focuses on still-image synthetic media and does not address real-time video deepfakes, which are less common but more difficult to detect.
-

9. Conclusion

AI-powered synthetic media is enabling large-scale fraud in Zambia, targeting vulnerable individuals through romance scams, fake product listings, and impersonation schemes. Existing detection tools are optimized for Western contexts and perform poorly on Zambian social media. Victims lack awareness of synthetic media threats and practical tools to protect themselves. Platforms and law enforcement are under-resourced to address the problem.

This research documents the problem in the Zambian context, develops a detection framework optimized for low-bandwidth mobile environments, and produces recommendations for users, platforms, regulators, and civil society. The findings confirm that synthetic media fraud is a serious and growing problem, but also that it is addressable through a combination of technical tools, user education, and institutional capacity-building.

The path forward requires action from multiple stakeholders:

- **Users** must become more critical consumers of digital information and build resilience against emotional manipulation

- **Platforms** must invest in local moderation, user education, and seller verification
- **Governments** must develop cybercrime capacity and support victim services
- **Civil society** must educate the public and support affected individuals
- **Researchers** must continue to study synthetic media fraud in African contexts and develop solutions tailored to local needs

Synthetic media fraud is not inevitable. With coordinated effort across these stakeholder groups, Zambia and other countries in Sub-Saharan Africa can build digital safety and resilience, allowing citizens to reap the benefits of internet connectivity while minimizing the risks of fraud.

References

Buchanan, T., & Whitty, M. T. (2020). Unfriending cybercriminals: Users' experience and perception of romance scams. *Cyberpsychology, Behavior, and Social Networking*, 24(4), 215–221.

Whitty, M. T., & Buchanan, T. (2012). Online grooming: The new playground for old tricks. *Journal of Psychology & Marketing*, 29(11), 860–872.

Interpol. (2025). 260 suspected scammers arrested in pan-African cybercrime operation. Retrieved from <https://www.interpol.int/>

Meta. (2023). Community standards enforcement report Q3 2023. Retrieved from <https://transparency.fb.com/>

Federal Trade Commission. (2024). Romance scams and investment fraud: 2024 report. Retrieved from <https://reportfraud.ftc.gov/>

DataReportal. (2024). Digital 2024: Zambia. Retrieved from <https://datareportal.com/reports/digital-2024-zambia>

Author Note

This research was conducted by John Kamfwa, a developer and security researcher based in Lusaka, Zambia, as an entry for the Apart Research Global South Hackathon.

Conflict of Interest

The author declares no conflicts of interest.