A Novel Framework Based on Deep Reinforcement Learning for Face Anti-Spoofing

Mr. Jangalapelli Shiva[1], Nadimetla Shiva Teja[2], Deekonda Prem Kumar[3], Talari Pavan[4]
[1,2,3,4]*Dept of Computer Science and Engineering (Data Science),*
CMR Technical Campus Hyderabad, Telangana

***Abstract*—There is a reduction in facial authentication due to emergence of other tricks of spoofing beginning with simple photo forgeries and progressing to very realistic deepfakes. CNNs are capable of performing a variety of tasks, but with some probability they create similar models and overlook novel types of fake images. I have observed that they gradually become ineffective as time passes. I created DRL-FAS to stop that. The former segment involves the reading of the overall image with a simple CNN. It determines the duration of standing, the place to search and the smallest details that are important. Such minor hints are fed into a RNN which learns through time. Lastly all these are compounded with general pointers in order to come up with a more powerful judgment. This will not stick to a given frame hence the model will continue to run throughout the video. We performed better in similar tests to those of Cai (in particular in against-untested attacks) in the group. The system identifies more deceptive fakes using less of the computing power.**

***Index Terms*—Face ID, machine learning, DRL-FAS, fraud (deception), trick of impersonation (spoof tricks).**

I. INTROUCTION

Nowadays, biometrics identification is everywhere in airports, smartphones, at the ATMs and even in security check up. A larger number of individuals use the same, and hackers devise better tricks to mislead systems. They attempt to exploit any kind of weakness.

The people initially were using basic materials such as printed pictures of faces that are ridiculous compared to the modern-day production of fake faces of high-quality using plastic and video replays that can even mislead human eyes in a moment. Deepfake video swaps began to work then, thus making people question the ability of cameras to determine the reality and the fake. Hackers bypass this type of anti-spoofing by making minute alterations to lighting and motion such that the ancient anti-spoofing devices overlook. I have observed that AI models gradually become very simple, such as the appearance of the eyes when tired due to repeated use of the image. Old fixes were reached using clear signals and yes/no checks but new attacks still defeat them or modify the environment in such a manner that the tools cease to work, and slow checks drag the entire system to a crawl when something goes wrong even slightly.

I have tried systems that do not always trust themselves and they are also slow to respond to new things as a time-warp that never needs to end. The earlier approaches which attempted to be initially clever resembled good enough but were then infiltrating old concepts anew causing a greater frequency of seeking out what was already familiar instead of genuinely merely changing anything. So, we switched to an alternate method of strong reinforcement learning - combining a high-level view with a keen eye, where somebody will quickly perceive a face and then scrutinize it, in case something is amiss. Thier basic CNN is scanning the entire image to identify shapes and general hints whereas a deep-reinforcement learning agent examines tiny sections of the image or work to identify strange light spots, jagged surfaces, weak shadows which appear unrealistic. The agent can choose when to zoom into a photo, jump forward or go into minute details that others familiar with the photo cannot see it can even hesitate. An RNN is used on every image, and it retains the variation in time and not in a single view only, therefore, it follows the appearance of mistakes. The experiments conducted in accordance with the research of Cai indicate that this system is better with the tricks unknown and it is responsive than imagined and thus it is capable of detecting fakes at an

optimal utilization of resources. However, that is not the only weird outcome of training that can be generated, i.e. the system may give too much attention to a peculiar place that makes no sense.

The new setup will be able to accommodate various types of inputs more effectively as compared to the old ones since they remain stable even when the lighting varies or which angles alter rapidly. The form remains minimal - less requirements, no continuous re-calibration - but it does enhance the manner in which models behave even on hard problems by integrating the broad knowledge with fast and fine-grained mending.

II. LITERATURE REVIEW / RELATED WORK

Author(s)& Year	Objectives	Techniques
Zhang, J., et al. (2024)	Hybrid CNN + ViT model for FAS.	Convolutional Vision Transformer (CVT).
Yu, Z., et al. (2023)	Efficiently adapt pre-trained ViTs for FAS.	ViT with a lightweight "S-Adapter" for fine-tuning.

Development of fake face detection has been gradual with researchers trying different methods to distinguish between the real and the fake ones. First, methods were based on manually made cues, including surface textures that were considered using tools, such as Local Binary Patterns (LBP) or analyzing assessments of fake prints, video playbacks, and masks found in old Personal Authentication Device (PAD) archives. Also, innate behaviors like blinking or displacing of the head had small protection against advanced methodologies of replay. Then followed deep learning which employed big image setups that also allowed models to look at finer details in both space; systems based on CelebA Things spoofing, together with more intelligent network architecture with central difference math and other training aids demonstrated that convolutional nets can out-pot deepwater feature chains; [1], [2], [3]. Extensions added pixel-wise labels and smooth CNN architectures- assisting in detecting the tiniest ailments of texture, and they made counter counteractions against feeble maneuvers more difficult [7]. Extended studies of presentation attack detection revealed performance fractures when stress-tested outside original condition demonstrated how most setups perform poorly with new tricks or unfamiliar tricks unless savvy adaptation mechanism intervenes [5]. During this period, reinforcement learning started gaining importance, supported by the meaningful results inspired by such methodologies as DQN, PPO, DDPG, and A3C in complex control problems. Progressively, the space shifted away because of formidable rule-based checks to more adaptable learning driven models that were capable of balancing the overall layout with details of suspicion.

III. SYSTEM DESIGN / METHODOLOGY

3.1 Existing System

Nowadays, biometric ID checks in case you are alive by detecting rapid eye movements or a slight movement of the head and the system attempts to sculpt out the intricacy of your face. The system fails however when a detailed bogus or a replicated image appears on a screen. What might be intelligent on tests disintegrates against intelligent fraudsters. Even posh high-end gadgets, which companies take pride to demonstrate, collapse very easily since their hastened inspections platform the nuanced hints that counterfeit faces reveal.

I have encountered such arrangements disregard minute details, which anyone keen would pick easily. Although scientists continue to release radical advances and brain models, even more basic machines continue to operate by old models which barely represent the change. And attach a polymer mask shaped adversely to resemble human face and even more high-grade depth sensors become lost.
It is like a clunky affair, all right. Canned shows find it difficult to contend with industrious sangwetters.

What's worse? These tools simply articulate real or fake as robots do not discuss the reasons why.
Meanwhile, there is research coming at a rapidly advancing pace, in systems that shift attention and are in operation, in systems that vary in response, or inquiring themselves whether something is amiss. However, a lot of live systems continue to utilize feeble hints, as though they have not kept pace. Both threats are becoming more dangerous and acute. The guards aren't keeping up. They are not perceiving a

great number of warning signs, and the time lag is increasing.

3.2 Proposed System

Our two-part facial authentication system is based on a pair of reactions that occur with the eye of a person when something does not look right (see Figure 3). It has two separate paths. The initial one is a regular CNN, like ResNet-18 that scans the entire face and transforms it into a small collection of objects. The task of this part is the primary one: shapes of faces, shadows, the texture of the skin, and other visual details which we are able to observe instantly. The second one involves the reinforcement-learned agent. This agent monitors the CNN feature map and determines where it wants to concentrate on and a particular area is chosen by the agent depending on the feature that is most critical.

The concept is based on DRL-FAS project by Cai et al. In the same work a network was trained to select helpful patches so that the system was able to recognize some fine details of a fake. Our sentinel examines the present circumstances, with CNN characteristics and an LSTM that accessible past and obtains subsequent steps in the image. It corrects its overall guess through an indicator of its correctness or incorrectness. Nevertheless, the more this is done over time it finds that certain details like the edges of photos, odd reflections, or blurry patterns tend to be indicators that someone is trying to lie to them. The movement of the agent is chaotic and may even be amusing in the initial runs but with subsequent training, it gains higher intelligence and precision in its movement.

Upon the completion of the training, the model also connects the general data and the final LSTM output and forwards it to a scoring layer. Practically, the CNN will capture the surface information, the LSTM will track the process of creation and the RL component will direct the point of focus. The various parts reinforce each other and the system as a whole looks at the face features in a manner that is similar to human beings detecting abnormalities.

3.3 System Architecture and Workflow

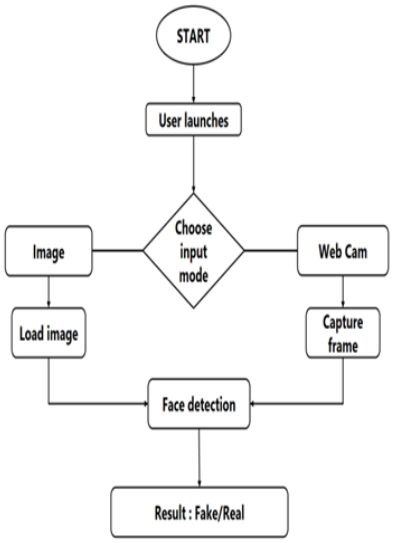

Figure 1: Architecture of A Novel Framework Based on Deep Reinforcement Learning for Face Anti-Spoofing

Figure 1 shows the full layout. One can follow the flow step-by-step and it is logical. It begins with a face picture or a live web cam picture. It is delivered to a CNN base typically the ResNet -18, which reduces it on a dense feature set. Then the process divides into two branches which operate parallel to each other.

The international direction is the first branch. It analyzes the entire feature map with the help of pooling or a couple of convolution layers and dense layers. The result is a universal feature array which incorporates the entire face, both look and lighting and general form.

Local reinforcement learning (RL) is utilized in the second branch. The RL policy does not look at the entire scene at once, but poses questions regarding the changing feature representation, and takes into account the hidden state of a recurring network. It in the next step selects the specific area that requires additional consideration.

Then it takes a small patch, such as 32×32 pixels, of either the feature map or the raw image depending on this option. The patch is inputted into an LSTM to process it. The LSTM changes its internal state with the help of every new patch. These rounds are repeated a certain number of times allowing the model to construct a view based on all the patches it has explored.

After the final stage, the patch-based features are merged with the global view using a fusion method. The resultant combination is passed to thick layers which determine whether the image is authentic or not. This is a good scheme since expansive structural data is mixed with detail tracking. The overall picture is provided by one path whereas subtle indications like borders, change in color, or even small flaws that are often overlooked by standard CNNs are identified by the RL-driven path.

3.4 Module Description and Functional Design

• CNN Backbone:

Our initial step is by using a pre-trained ResNet, more commonly ResNet-18, and we need not to learn a new model. It transforms every face image into a set of numbers- some 512 values on a 7×7 grid or the likes. Then the information is divided into two branches.

• Global Features:

The global part stays simple. In place of detailing the various steps, we simply average the entire feature map or just do a 1x 1 convolution, after which we submit the outcome to a dense representation that collects all the values into a single, fixed identified metric. This is the direct entry into the last prediction stage.

• RL Agent (Policy Network):

This was obtained by merging the CNN output map and the final RNN step by the agent. These are pumped into a small tight net which makes the decision of where to search. Network results generate a probability of each site in the grid where the patch may be. Accepting these probabilities, it chooses an (x, y) coordinate during the training and the location with the greatest probability during operation. The total number of locations is 49 with a 5127x7 feature layout.

• Patch Extraction:

Once the agent has chosen a spot we take a patch around it, of a small size -3 in feature-map terms. The local feature extractor is then fed on this patch.

• Local Feature Extractor (RNN):

We use an LSTM here. Every time an additional patch is inputted, LSTM changes its hidden state and, progressively gathering all the information on number of patches that it has received up to that time. The concealed condition begins as all the zeros. Once T has circled the last hidden state the entire local feature summary is a result.

• Fusion and Classification:

The final local (global) vector is added with the global feature one and they are subjected to several dense layers. A softmax will then generate the likelihood that the image is real or fake. Cross-entropy loss is applied to direct the learning in training.

Training:

This entails modifying its environment to offer an appropriate reinforcement schedule (including positive, negative, or neutral reinforcement), which is anticipated to maintain a specific aspect of the behavior (Wolff, 2009), either via alternative means or by utilizing a shaping process to alter the learning environment (Millar, 2003). <|human|>Training: This involves modifying its environment to provide a proper schedule of reinforcement (positive, negative or neutral), that is expected to perpetuate a particular aspect of the behavior (Wolff, 20

The policy network is trained through a basic RL algorithm e.g. REINFORCE. Each time the agent is provided with an image, it has a reward, usually +1 in the event of a correct choice, and 0 in the event of a misjudgment. This reward is multiplied with the log probabilities of the chosen patches and the model transforms its weights according to the signal that it obtains. Takeing numerous steps the system slowly adapts to which patch order is more helpful and pays more attention to the productive regions.

IV. IMPLEMENTATION

4.1 Algorithms used

• Convolutional Neural Network (CNN):

Our first choice is a basic CNN which is ResNet -18, easy to install and works. It acquires overall characteristics of faces thus like having a glance at a picture. It is typically trained in advance on ImageNet, and AE is then trained to learn faster on fake-amples. To compute the weights, we have back-propagation, nothing special, cross- entropy running the step.

• Recurrent Neural Network (RNN):

The LSTM processes the sequences in a real time manner. Every time through, it reads a small chunk

and all the information it has previously read. Then it incorporates the new information into its long-term memory. The final state provides a slight overview of all the places that the model searched. It is trained along with the rest therefore the entire chain is trained as a whole rather than being trained in parts.

• Policy Network (RL Agent):
This one makes the decisions. It examines the existing information, the CNN output along with the memory of LSTM, and selects probable locations of patch centres. It is a simple network as it consists of only dense layers with softmax on top of. Nevertheless, it guides the next movement of the gaze. Learning is an experiment that seeks out in all directions; testing is one that is immediately driven to the first options.

• Reinforcement Learning (Policy Gradient):
Teaching the decision-maker? Not tough. Feedback is obtained on every attempt: there is a nod of approval in case of success, no no in case of failure. It starts to prefer actions which provoked victories previously. Add a baseline to the system, to hold the changes in the system reasonable, otherwise the changes become wild, like morning coffee. All in all, it operates in a similar way as classic REINFORCE but is not as messy.

• Feature Fusion and Classifier:
When we have a global picture and the comprehensive local picture they are unified into one embedding. The overall representation undergoes numerous transformations and finally, a softmax is used to determine whether it is real or fake. Cross-entropy is used to measure the difference between predictions and the truth, and we pass any discrepancy on through each part of the network the classifier, the LSTM and the convolutional network.

• Training Procedure:
The input in each mini-batch is provided in two different categories simultaneously. Labeled data ensures that the main model is on track, the CNN, LSTM and classifier are tightened so they do not drift and that it is not far away wrong. The other input is less rigid and is almost a test. This policy mechanism attempts numerous actions, receives rewards and nothing and repeats its patterns on the basis of which actions were effective. There are numerous rounds, and the entire system begins to work more effectively. It is just a matter of looking: a watcher notices the difference between the results shown by the feature detector: it becomes more accurate, the patch selection does not shift a la Price Anarchie anymore, the parts gradually assemble into a concrete rhythm which was not present initially. It might appear disorganized initially but eventually the components are in perfect and natural order.

V. RESULTS AND DISCUSSION

The primary screen has got a prominent entry point. It loads automatically and rests until the time you need it. Everything is self-evident at first glance, including preview area, alert icons, trust meter, etc. After being configured, the tool is ready to either process photos or stream of camera depending on your preference. A change of mode to real-time demonstrates the instantaneous behavior of the model in regard to the motion. Scanning of each frame of the camera is immediately, without delay. Once the real face is exposed, then recognition begins fast and confidence is stable. Whenever the lighting is altered or the individual swivel his head, the system is instantly recalculated in a smooth and speedy manner. Its view and playback is an indication of the seamlessness of the CNN, LSTM, and RL parts. It is followed by the fake-detection test with an ID photo. The system lands into action whenever a flat or a plastic face passes before it. It scans irregular edges, weird shadows or rough faces, characteristic of the patch selection tool which was educated on trial. Minor touches are pointed out and it is not a mere guess that it is not a real picture. It is not only a sound alarm, but the areas that are highlighted indicate which items were switched-off. The same occurs in case of an artificial face on screen. Its effect of light pattern appears as an abomination of light giving the surface an artificial lustre, weird micro-reflections, or bizarre shadows. It is instantly pronounced a fake by the system. It does not hang when the highlights are overexpanded as was the case with old systems. It continues working automatically, only when there is something new it draws attention to. The outcome is quick and dependable making the decision remain unchanged even when the forgery is meticulously prepared. All these characteristics indicate that the system is smooth-running. It performs effectively with

authenticated users, it can learn to identify counterfeit threats at a very fast rate, and it remains stable when the environment is altered. Such a mixture of general awareness and flexible close-range inspections is in its merits immediately when you are viewing them in the actual life.

VI. CONCLUSION AND FUTURE SCOPE

The present project developed a deep reinforcement learning system to track counterfeit faces. It works with a wide-range convolutional neural network and an attention mechanism which functions in a manner that is similar to the way individuals gaze at faces. The model does not see everything at once as looking at the whole picture and then the odd details which are actually zoomed. It was aimed at the network to do the same, that is, to survey widely, and then to pay attention to significant details. It worked well.

The system was also highly precise without additional complexity on common test data, and was performing as well as the best recent methods. Once again, it examines the entire picture first and then pays attention to peculiar places. That was how the network was trained and it succeeded. At that, the broad context and the close detail elements were brought together to allow information sharing rather than getting in the way. When this was integrated, the entire system was easily more effective than single shot techniques.

There are still numerous issues. The model can break down on inputs never encountered before and this is something that has frequently been ignored. In order to operate successfully in the actual world, these systems have to be able to manage unfamiliar threats in a relaxed manner. Such strategies like domain adaptation or learning without original samples are also promising; it will be interesting to observe what happens to them. This may be complemented by additional signals, including depth maps, infrared, or any other feature that is built by the camera. There is a split-up arrangement in which information is enlisted to the gadgets, in order to ensure the security of your privacy.

Indeed, by making the system lighter phones and other small devices can run it without heating or being slow. Weird concepts can be of use, such as ingenious mechanisms that sneak a peep without being told what to look at, or devices that tell when some strange face is around even when it cannot be compared with the victims of those threats. Megabits, megabytes of data are important- lumpy lighting, grey shadows and other natural environments present challenges to the system. Sealing these gaps would enable these devices to move out of the laboratory and operate successfully on crowded streets, in shady areas or during nighttime.

REFERENCES

[1] Y. Zhang, Z. Yin, Y. Li, G. Yin, J. Yan, J. Shao, and Z. Liu, “CelebA-Spoof: Large-scale face anti-spoofing dataset with rich annotations,” in *Proc. European Conference on Computer Vision (ECCV)*, 2020, pp. 70–85.

[2] Z. Yu, C. Zhao, Z. Wang, Y. Qin, Z. Su, X. Li, F. Zhou, and G. Zhao, “Searching central difference convolutional networks for face anti-spoofing,” in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 5295–5305.

[3] Y. Liu, A. Jourabloo, and X. Liu, “Learning deep models for face anti-spoofing: Binary or auxiliary supervision,” in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 389–398.

[4] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li, “A face anti-spoofing database with diverse attacks,” in *Proc. International Conference on Biometrics (ICB)*, 2012, pp. 26–31.

[5] S. Jia, G. Guo, Z. Xu, and Q. Liu, “A study on presentation attack detection in face recognition,” *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1091–1104, 2020.

[6] Chingovska, A. Anjos, and S. Marcel, “On the effectiveness of local binary patterns in face anti-spoofing,” in *Proc. BIOSIG Conference on Biometrics and Security*, 2012, pp. 1–7.

[7] George and S. Marcel, “Deep pixel-wise binary supervision for face presentation attack detection,” in *Proc. International Conference on Biometrics (ICB)*, 2019, pp. 1–8.