Facial Recognition in War Contexts: Mass Surveillance and Mass Atrocity

*Juan Espindola**

As the use of facial recognition technology (FRT) in the policing activities of contemporary societies gains ground, the criticisms directed at it grow concomitantly. The deployment of FRT in authoritarian and liberal democratic regimes alike to persecute ethnic groups, repress political dissidents, or conduct widespread unjustified surveillance—particularly when the technology is integrated into closed-circuit television, or CCTV, systems—has been aptly described as a political and social menace. Even when relied upon for legitimate purposes, FRT has come under fire for its insidious biases, which disproportionately hurt minorities.

FRT as a form of intelligence has recently made a prominent public appearance in the theater of war. During the early months of Russia's invasion of Ukraine, Ukrainian authorities relied on FRT as part of the country's defensive activities, harnessing the technology for a variety of purposes. FRT has been deployed to unveil covert Russian agents operating amid the Ukrainian population; to reveal the identity of Russian soldiers who committed war crimes; and to identify dead Russian soldiers. This constellation of uses of FRT—in a war increasingly waged on the digital and information front—raises significant concerns and

Juan Espindola, National Autonomous University of Mexico, Mexico City, Mexico (juanespindola@comunidad.unam.mx)

*I would like to thank Cécile Fabre, Ross Bellaby, Ron Dudai, Alex Leveringhaus, and Rhiannon Neilsen for their feedback on an earlier version of this draft. I am particularly indebted to Fabre for her detailed written comments on the piece. Finally, I would like to thank the editorial team of *Ethics & International Affairs* for their careful and incisive revisions of the piece. This work was supported by grant UNAM-PAPIIT IA400523.

Ethics & International Affairs, 37, no. 2 (2023), pp. 177–192.

doi:10.1017/S0892679423000151

therefore warrants ethical examination. Some of these concerns parallel those to be found in nonbellicose contexts; others, however, are unique to war-torn settings.

To explore the ethical issues surrounding FRT as an aspect of wartime intelligence, this essay begins by building the best-possible case for FRT as a justifiable form of intelligence, based on the claim that it can be instrumental in deterring and preventing threats to individual rights and in fulfilling humanitarian duties toward combatants and their relatives, especially in situations of mass atrocity. The essay then highlights some of the serious concerns with FRT, including the infringement of informational privacy; the indiscriminate and disproportionate harms it may inflict, particularly when the technology is coupled with social media intelligence; and the potential abuse of the technology once the fog of war dissipates.

Many of both the justifications and concerns about FRT that I discuss here are the subject of Cécile Fabre's recent book *Spying through a Glass Darkly: The Ethics of Espionage and Counter-Intelligence*,[1] which offers the most systematic and rigorous defense of espionage and counterintelligence as a permissible, even mandatory, form of self-defense in the face of threats to fundamental rights. Other justifications and concerns can be built on the basis of Fabre's account. The essay will concentrate on the Russian invasion of Ukraine for expository purposes, but its implications may be extrapolated to other conflicts, as the technology is increasingly used in a variety of settings.

Before I address these issues, it is important to define what is meant by FRT in the context of this essay. FRT is a form of artificial intelligence involving "the automated extraction, digitization and comparison of spatial and geometric distribution of facial features to identify individuals."[2] It comes in four varieties based on their function, each with its own ethical implications. These functions are, from basic to more complex: "detection," which does not collect personally identifiable information; "characterization," which collects information such as gender, age range, and emotional indicators, but does not collect or retain personally identifiable facial templates; "verification," which does collect and retain facial templates, and then uses a one-to-one matching system in which software determines whether the person is who she claims to be (for example, a smartphone scans your face to determine whether it matches the saved template); and, of main relevance for the purpose of this essay, "identification," which also creates an identifiable template of a unique person but uses it in a one-to-many

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press

matching system, where the system compares a collected image against an existing database.[3]

Preventing Threats, Deterring War Crimes, and Discharging Humanitarian Duties

Ukrainian authorities and allied nonstate actors engaged in numerous acts of espionage prior to the Russian invasion. Most notably, they conducted open-source intelligence, which helped Ukraine detect Russia's military buildup around its frontiers and anticipate the invasion, notwithstanding Russia's protestations that it would refrain from invading. Russian military mobilization was revealed by a combination of satellite pictures, including some captured by private companies, and videos and images from social media platforms, particularly through TikTok, coming out of Russia.[4] Ukrainian authorities and the country's allies conducted further acts of espionage once the invasion began, some of which involved the joint use of biometric and social media intelligence. In particular, they deployed FRT to identify enemy operatives (spies, combatants) and fallen soldiers on both sides of the trenches. Clearview AI, a company whose facial recognition service is under intense legal scrutiny in the United States and elsewhere, spearheaded these efforts. Several agencies received access to the services of Clearview AI, including the national police of Ukraine.[5]

Why should the use of FRT in this context be considered an instance of espionage or counterintelligence? I argue there are three key reasons: First, FRT unveils the identity of combatants by combining biometric data with the collection of vast amounts of personal information from social media platforms and other venues on the Internet. The resulting intelligence amounts to mass surveillance.[6] Second, the use of FRT in this context is a case of *defensive* counterintelligence because it involves intelligence work conducted against a foreign power as a response to an aggression. Third, the information thereby revealed—the identity of Russian soldiers—is of the type that Russian officials would prefer to keep secret since its dissemination can be (and has been) used with the prospect of exerting pressure on Russia's civil society to speak out and mobilize against the war and on international civil society to condemn the war. This suggests that the goals of this intervention oscillate between counterintelligence and warfare, a point to which I return below.

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press

Is the use of this technology as an intelligence tool morally justified? Just as with the open-source intelligence conducted prior to the invasion, Ukrainian authorities have sufficient justification to deploy a wide range of intelligence interventions to identify Russian operatives. Many of these justifications are consistent with Fabre's normative case for espionage. Her case rests on two foundations. The first foundation is the protection of fundamental rights (understood as protections for the possession and enjoyment of certain capabilities) as well as the duty of individuals to protect each other from violations of these rights. On Fabre's account, protecting fundamental rights is for the most part understood preventively: rights are protected by parrying threats to them. Thus, espionage is morally justified (even mandatory), but "only as a means to protect oneself and third parties from violations of fundamental moral rights or risks thereof in the context of foreign policy writ large."[7] The second foundation in Fabre's case for espionage is an account of when individuals are justified in harming one another to defend these rights; that is, when they are liable to self-defensive harm. Individuals may be harmed in self-defense if they contribute to a morally unjustified threat that significantly affects these fundamental rights, or if they wrongfully fail to do their part in thwarting this threat.[8]

Based on this framework, Fabre articulates a defense of a pro tanto permission and duty to spy as a means to thwart rights violations. In her view, nation *G* is permitted to spy on nation *B* if *B* mounts an unjustified attack on *G*; *G*'s espionage must also be understood as part of its defensive action, which, if it meets the criteria of effectiveness, necessity, and proportionality, is justified. *G*'s espionage serves to assess *B*'s war-fighting capacity, an assessment that would enable *G*'s forces to tailor its defensive actions. Note that spying is also permitted in response to other scenarios, not just as a response to a bellicose threat. It may be a response to rights-undermining foreign policy in bilateral or multilateral scenarios. Unwarranted espionage on the part of *B* may also be just cause for its target, *G*, to engage in defensive counterintelligence.

Fabre's theoretical framework can help us elaborate plausible justifications, as well as potential objections, for the use of FRT as an intelligence tool during wartime. Consider in greater detail the three uses of FRT by Ukrainian authorities in the current conflict. First, on the basis of preventing threats, one can justify Ukraine's use of FRT to unveil Russian infiltrators amid the Ukrainian citizenry. Consider the example of L'viv, a city on the border with Poland and a transit point for refugees. As a result of the refugee crisis, its population surged by about four

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press

hundred thousand people during just the first month of the war. During that time, Ukrainian intelligence dismantled twenty Russian sabotage groups and arrested 350 suspected saboteurs there, thanks in part to the use of FRT. Uncovering these *dyversanti*, or saboteurs who mix with the Ukrainian population to sow mistrust and alert Russia about potential targets,[9] would seem to provide a compelling reason to use Clearview's mobile app to scan faces at checkpoints or while out on patrol.

Another use of FRT that can be justified is tied to Ukraine's pursuit of a different goal: identifying Russian soldiers who commit war crimes, such as the killing of innocent civilians and looting in the occupied territories.[10] The aspiration is that FRT can help identify the perpetrators. Once collected, the evidence can be the basis for later prosecution at tribunals such as the International Criminal Court, with the hope that the fear of prosecution will provide deterrent effects, albeit admittedly modest ones.[11] To illustrate, as was amply documented during the first months of the invasion, Russian combatants committed war crimes systematically, the Bucha massacre being one of the most egregious examples. Owing in part to FRT, as well as social media intelligence and open-source intelligence, several members of a battalion of the Russian army were identified as the main perpetrators.[12] While it is true that this kind of future punishment–based deterrence is not the same as deterrence to prevent an imminent threat, as Fabre envisions it, both approaches are ultimately conducive to protecting individual rights.

The third and most controversial use of FRT that Ukrainian government agencies have deployed is related to yet another goal: identifying dead combatants. According to reports of the early stages of the conflict, authorities in Ukraine photographed the faces of dead belligerent soldiers, matched the resulting images with the databases provided by FRT companies, and then proceeded to disseminate the images (and the matches) over social media. This usage of FRT cannot be justified on threat-prevention grounds, as were the previous two; instead, the justification must be founded on an entirely different basis altogether.

Such a foundation can be, I think, grounded in the fulfillment of humanitarian duties, in particular toward fallen combatants, including enemy combatants and their relatives. During any war, belligerent parties may lack the means (or the willingness) to promptly identify the bodies of the fallen and to allow the enemy to collect the bodies of its citizens. FRT may help overcome this challenge because of its ability to rapidly identify deceased soldiers. Along these lines, Ukraine's Ministry of Digital Transformation justified the use of FRT (and the subsequent

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press

dissemination of images of deceased soldiers over social media) as "a courtesy to the mothers of those soldiers . . . to at least let families know that they've lost their sons and to then enable them to come to collect their bodies."[13] As part of this campaign, Ukrainian officials even asked parents to send in their own DNA samples to help determine whether their sons had been killed in combat.

On the face of it, this motivation puts facial recognition on solid legal and moral grounding. Legally speaking, the Geneva Conventions impose the obligation to facilitate the return of the bodies of dead soldiers to the home country upon the country's request or that of relatives.[14] Morally speaking, regardless of whether fallen soldiers are just or unjust combatants or civilians, their relatives and friends have a right to retrieve their bodies. Belligerent parties, in turn, have a humanitarian duty to enable the retrieval as promptly and effectively as possible, to the extent of their abilities.[15] In any case, handing over the bodies of the deceased to their relatives allows the latter to adequately mourn the former and spares them from experiencing what Pauline Boss calls "ambiguous loss."[16] Ambiguous loss occurs when the loss of a family member or loved one is enveloped in uncertainty: she or he is physically absent but psychologically present. Ambiguous loss leads to unresolved grief, freezes the mourning process, and contributes to the deterioration of the mental health of the bereaved; these factors are not present, or not in the same magnitude, in cases not involving ambiguous loss. Some of the effects of ambiguous loss include feelings of being alone or detached from other people, and confusion about one's role in life or a diminished sense of one's identity, all of which undercut people's capabilities in Martha Nussbaum's sense of the term (their ability to engage in meaningful social relationships and to frame a conception of the good life),[17] and thereby undermine their fundamental rights. If FRT can be instrumental in helping individuals overcome ambiguous loss by accelerating or assisting in identifying dead bodies, then it is instrumental in protecting fundamental rights and is therefore prima facie justified.

Objections to FRT as a Justifiable Form of Counterintelligence

Suppose we agree to view Ukraine's use of FRT as a counterintelligence activity pursuing justifiable goals, such as preventing threats, deterring war crimes, or overcoming ambiguous loss. Notwithstanding these contributions, what are some objections to FRT in war contexts? I turn now to examine three of these

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press

objections and assess their normative purchase: the privacy objection; the slippery slope objection; and the warfare objection. The first two objections track, with some adjustments, those leveled against the technology in nations that are not engulfed in war. The final objection addresses concerns that arise exclusively in circumstances of war. I shall argue that the privacy objection is defeasible but that the other two objections are a genuine source of concern, which renders the use of FRT in war contexts impermissible.

The first ethical concern is that FRT may infringe upon the interest in informational privacy. Because it draws on vast amounts of data (specifically, images), which it scrapes from social media platforms and other Internet outlets usually without the consent of users, FRT involves a form of mass surveillance that violates the informational privacy of such users. Informational privacy allows individuals to control access to their information, including their facial images, excluding other individuals or organizations such as the state from such access if so chosen. Informational privacy is valuable because it is closely connected to the fundamental value of autonomy—broadly, the liberty to think and do as one chooses. The implications of the infringement of information privacy are significant: such infringements can impede the pursuit of personal projects by interfering with the plans individuals craft for their lives. Collectively, they can undermine liberal democracy, as when citizens no longer feel at ease demonstrating or speaking out against authorities for fear of reprisals enabled by government surveillance through FRT.[18]

Ukraine's technological feat with FRT has been accomplished precisely because the services of companies like PimEyes, FindClone, or, most controversial of all, Clearview AI violate informational privacy. Clearview AI claims to have two billion images from the Russian social media service VKontakte at its disposal, which its technology draws on to identify Russian faces.[19] The privacy objection contends, then, that Russians who posted their data (pictures) on this social media platform did not consent to have them collected and stored by Clearview AI, let alone by foreign agencies, which acted wrongly by collecting and storing them, nonetheless.

We should not minimize the risk that data collection and storage pose for the privacy of individuals. Yet Fabre's account of espionage can help us assess the normative weight of this concern. As she convincingly claims, the privacy objection to mass surveillance may be defeated in contexts where privacy infringements can decisively contribute to discharging a duty of protection, understood as the timely

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press

detection of threats to fundamental moral rights. The critical step is to weigh the magnitude of the harm of placing an individual or group of individuals under surveillance, the magnitude of the harm that such an action might forestall, and the magnitude of the harms relative to each other.[20] When we turn to the war in Ukraine, a plausible case can be made that the magnitude of the harm of infringing on the privacy of Russian users of social media platforms, while significant, is moderate compared to the magnitude of harms that such an action can forestall. A privacy breach, in this circumscribed context, may therefore be permissible.

A different argument that can be made to justify an infringement on the privacy rights of Russian social media users is that they owe this partial forfeiture of their privacy rights as a form of compensation for the wrongful harms caused by the regime that waged war in their name. This case is more easily made with respect to citizens whom we can describe as "wrongful beneficiaries" in Avia Pasternak's definition of the term: they know about the source of the benefits but nonetheless desire them and actively pursue them, even when they would be entirely free to forgo the benefits. Those who profit from the markets for stolen goods that solidify as the spoils of war make their way into Russia are wrongful beneficiaries of the war in this sense. Wrongful beneficiaries, as Pasternak rightly claims, have weighty compensatory duties to victims arising from their indirect contribution to the wrongdoing, and it could be contended that forfeiting their right to informational privacy is a manner of compensating victims, even if a very imperfect manner.[21] Even Russian citizens who are not wrongful beneficiaries may have to forfeit part of their privacy rights in order to aid efforts to end the war. This is for the same reasons that ordinary Russians must bear the costs of the imposition of economic sanctions on Russia (with the caveat that they should not lead to excessive hardship): while they may not benefit from the war, they contribute to it by funding it through their taxes, by supporting the soldiers that wage the unjust war, and so on. Admittedly, many of them have no choice but to do these things. Most of them can emigrate and protest only at an unbearable cost. Nonetheless, they are implicitly sustaining the war effort. So, as Pasternak and Stemplowska argue, "Perhaps they too can be expected to incur some harm if such harm could stop the war."[22] At any rate, all these reasons converge to provide a plausible normative basis for the infringement on the informational privacy of Russian social media users.

Let us turn to a more powerful objection to FRT—the "slippery slope objection," as Evan Selinger and Brenda Leong call it.[23] A slippery slope objection is

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press

an argument against an action, *A*, that is not in itself objectionable but whose performance will lead to the performance of a chain of actions that in the end will lead to action *Z*, which is objectionable. Or put differently, in a slippery slope argument, "it is not permitting the instant case that worries us, but rather the possibility that permitting the instant case will lead to the danger case."[24] In the case of FRT, the concern is that encouraging or tolerating its use for permissible purposes in war contexts, such as identifying enemy agents or fallen soldiers, might open the door to objectionable goals in a postbellum scenario, such as its deployment to conduct unjustified intelligence against innocent nationals or its use by local or transnational nonstate actors whose goals, while not necessarily unjust, are opaque to us. I shall focus on these two potential cases of so-called "function or scope creep,"[25] although some authors envision worse end-scenarios for FRT, such as its incorporation into systems that use automated decision-making to direct lethal force (autonomous killing machines).[26]

Some slippery slope arguments are fallacious in that they identify neither the causal steps that take us from innocuous and unobjectionable actions or scenarios to the morally undesirable ones, nor the likelihood that such undesirable scenarios will materialize. By contrast, reasonable versions of slippery slope arguments "explicitly specify a plausible mechanism that could drive slippage from one step to another" and "rigorously explain why the mechanisms deserve due consideration."[27] The mechanisms that may trigger the causal chain that takes us from the unobjectionable to the objectionable case are, for example, "cost-lowering" and "attitude-altering," or the creation of the political momentum for doing so.[28]

What are the mechanisms that could drive us down the slippery slope in the case of FRT in Ukraine or similar contexts? According to one human rights activist, companies promoting FRT "are eager to exploit the humanitarian crisis in Ukraine to normalize the use of their harmful and invasive software."[29] This criticism describes the attitude-altering mechanism of normalization, identified by Selinger and Leong as used in non-bellicose settings. In peaceful times, normalization arises from the positive representation of a technology in the media (portrayed as fun and time saving). Once citizens are habituated to the joyful and "efficient" environments it creates, people are more inclined to accept its usage in an increasing number of realms in social life. CCTV cameras are an example of this kind of pernicious habituation: the enduring loss of privacy they entail is perceived to be less troublesome if they are deemed to be effective in preventing crime, and with the passing of time citizens will have lost sight of the trade-off.

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press

In the context of war, it is hard to think of a better reputational boost to a technology than the appreciation of its contribution to countering an unjust war of aggression, and such a contribution is precisely what FRT may provide in Ukraine today. The worry, then, is that once a positive image for the technology is entrenched in wartime, it will be difficult to dislodge that technology in the aftermath of war. The worry is reinforced by the fact that citizens amid war are relatively powerless to resist the use of the technology, and such powerlessness may breed an adaptive preference for it.[30] It might be countered that this concern is overblown because citizens can discern the difference between what is acceptable in war and what is acceptable in peacetime. The whole point of resistance is to make and accept sacrifices that they would not otherwise tolerate in ordinary circumstances. Consider, however, that after war people may be inclined to discount the risks of technology and, by contrast, inflate those risks associated with the presumed activities of the invader, analogous to cases where citizens are willing to let domestic law enforcement authorities acquire advanced military technology when they have a distorted sense of the security threats they face.

Another causal driver of the slippery slope alluded to by human right activists is the likely reluctance of Ukrainian authorities to "hand back" FRT to Clearview AI once the conflict is over and instead proceed to use the technology for illegitimate purposes, such as surveilling Ukrainian citizens without proper justification.[31] The reluctance is of concern because of an institutional consideration standing in the background: the absence of safeguards to curtail potential abuse of FRT after the war is over. In his discussion of the collection of metadata for intelligence purposes, Michael Skerker argues that in states with robust rule of law and high ethical public service standards, the risks of metadata abuse are likely to be lower than those of a terrorist attack or a military or intelligence operation facilitated through contact with a local actor. The opposite is true when these institutional qualities are lacking.[32] The same logic holds with respect to FRT: Whether the risks associated with it are realized or not, and the extent to which they are realized, depends on the presence or absence of these institutional qualities. In the case of Ukraine, the institutional capabilities that may keep FRT in check in a postwar scenario may be absent.[33] Furthermore, it is precisely in the wake of a destructive war that, for obvious reasons, institutions are at their weakest, particularly those tasked with functions beyond basic security and reconstruction.

It could be argued that FRT can be programmed to contain technological safeguards to eliminate or reduce potential for government abuse. Clearview AI can

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press

vet the technology to impose constraints on its use, and to make sure only authorized government officials have access to it. It can put in place relatively simple safeguards such as two-factor authentication, such that at the very least public officials from branches of government that are not tasked with security functions do not have access to the technology. Moreover, since the technology is operated through a cloud service, Clearview AI retains the ability to revoke access in cases of abuse of the technology—and would do so, according to its CEO, in cases of egregious abuse only.[34] However, these technological solutions to potential government abuse are inadequate, not only because, as the case of the NSO Group (creator of the Pegasus spyware) shows, corporate assurances are weak at best, if not fully unreliable. Most importantly, trusting corporate assurances simply displaces the concern for barriers against abuse from the public to the private sector. The question then becomes whether there are institutional guarantees in place to prevent third parties with profit aims such as data brokers from, say, reusing or selling personal data for their benefit.[35] Shifting the responsibility to curb abuse from the public to the private sector might even aggravate the risks associated with FRT. For one, society's leverage to constrain corporate actors, particularly technology companies, may be weaker than the tools to rein in government agents, as the rise of Google and Facebook shows.[36] For another, by leaving it up to corporate actors to make determinations over what constitutes an egregious use of the technology, citizens would end up vesting corporate actors, which lack political legitimacy, with the authority to make decisions that affect their lives. Technological corporations would come to set common standards for digital governance, a task that ought to be a matter of collective decision-making. This creates a legitimacy deficit.[37]

The final objection to the most controversial use of FRT relates to the act of identifying the bodies of dead soldiers and disseminating their images on social media. As was mentioned before, Ukrainian authorities justified the strategy on the grounds that it fulfilled a humanitarian duty. Several objections speak against the practice, however. First, the technology can make mistakes. These do not necessarily derive, as in ordinary contexts, from racial or ethnic biases incorporated into the technology (with the corresponding worries about unjustifiable discrimination) but rather from the crude fact of postmortem decomposition. As one might expect, the efficacy of FRT is greatly impacted by the degree to which decomposition has occurred.[38] This is a particularly relevant concern in the context of war, where bodies may be subject not only to advanced decomposition but

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press

to disfigurement or mutilation. When postmortem misidentifications rise beyond a reasonable threshold, the high number of false positives can undermine the humanitarian duties FRT is supposed to assist.

Furthermore, the mere act of disseminating the images of fallen soldiers, regardless of whether their identity has been revealed through FRT, might be construed as disrespecting the dignity of the dead, thereby coming into tension with the Geneva Conventions. Admittedly, the Conventions only make sparse references to the treatment of the dead. One of its provisions calls on belligerent parties "to protect [the killed and wounded] against . . . ill-treatment" when "military considerations allow,"[39] and another provision calls on parties to respect "the remains of persons who have died for reasons related to occupation or in detention resulting from occupation or hostilities."[40] A broad interpretation of these provisions could equate the dissemination of images of dead soldiers with a degrading treatment in view of the widespread expectation that the public display of corpses for purposes other than ceremonial ones is, with some exceptions, undignified.[41] This is a controversial position insofar as one could easily call into question whether dead people have an interest in defending themselves against undignified treatment since they cannot experience the humiliation that accompanies it and therefore lack the capacity to be affected by it. Even then, there is the experience of relatives, who by contrast may be affected by it, to reckon with.

The most serious problem with the practice, however, is that, by admission of Ukrainian officials, it is not carried out solely, and perhaps not even principally, to fulfill humanitarian duties; it is also as a means of self-defense. The point of disseminating the images of dead Russian soldiers (again, with or without an FRT identification appended to it) is to document the realities of the battle zone, to counter false claims by Russia about the nature and the human toll of its military operations, and to seed distress and anger in Russian families in the hopes of galvanizing opposition within Russia to the invasion. That this is a central goal of the strategy is made clear by the fact that it is embedded within a broader campaign of public awareness of the horrors of war. To this end, accounts set up by Ukrainian authorities and sympathizers on platforms such as Telegram, Twitter, and YouTube streamed extremely graphic images of the war, inviting civilians behind enemy lines to confront them.[42]

Putting aside the goal of raising awareness, the problem with the strategy of distressing and angering the relatives of deceased soldiers is that it amounts to a form of psychological warfare, in flagrant violation of the Geneva Conventions and of

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press

jus in bello principles, more generally.[43] The purpose of the conventions and of the more abstract principles of just war and just intelligence theories is to protect civilians from the ravages of war by clearly circumscribing what is and what is not permissible self-defense, thereby sparing them from unjustified harm. Under just war theory, liability is critical to justify the infliction of harm on others: only those who have forfeited their right not to be attacked (for example, because they conducted an unjustified aggression) are liable to be harmed. Self-defense acts that are not bound in this way are indiscriminate.

Seen through the lens of permissible self-defensive harm, the use of FRT in the manner under discussion may be too blunt of an instrument of warfare because it inflicts indiscriminate harm.[44] Along these lines, Russian families are not liable to be subjected to the distress raised by FRT for the sake of bringing the war to an end when they did little, if anything, to encourage the war. When they are so targeted, their suffering is used as a means to mobilize them and other Russian citizens against the war. They are targeted *opportunistically*. Contrast this to the rationale for targeting Russian soldiers: they are targeted *eliminitavely*, as Seth Lazar puts it. The goal of their suffering is not to derive a benefit that could not be otherwise obtained; rather, they are targeted "to solve a problem that they themselves pose."[45] Admittedly, as I mentioned earlier, Russian citizens in general may have to tolerate that some of their rights, like the right to informational privacy, are curtailed as a contribution to potentially ending the war or accelerating its end. But it is one thing to accept the curtailment of a right and quite another to be made the target of (psychological) aggression.

Conclusion

The use of FRT, coupled with social media intelligence, as a weapon of warfare and a resource to conduct intelligence is on the rise. There is a plausible case to make about the permissibility of its deployment both to acquire information to prevent harm and to fulfill humanitarian obligations in certain contexts. A salient objection against its use is that it violates the privacy of some social media users, but this objection is not compelling. More convincing are the objections that FRT may be weaponized, inflicting indiscriminate harm to Russian citizens, and that since institutional safeguards against abuse of the technology tend to be weak in the aftermath of war, the use of FRT during wartime could embolden its postbellum use in Ukraine and thereby undermine, at a later time,

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press

the rights of those citizens it is supposed to protect at the present time. In the final analysis, whether the wartime benefits of FRT outweigh its postbellum risks is a matter to be decided contextually.

Notes

[1] Cécile Fabre, *Spying through a Glass Darkly: The Ethics of Espionage and Counter-Intelligence* (New York: Oxford University Press, 2022).
[2] Evan Selinger and Brenda Leong, "The Ethics of Facial Recognition Technology," in Carissa Véliz (ed.), *The Oxford Handbook of Digital Ethics* (Oxford: Oxford University Press, 2021), p. 31.
[3] Ibid. See also Denise Almeida, Konstantin Shmarko, and Elizabeth Lomas, "The Ethics of Facial Recognition Technologies, Surveillance, and Accountability in an Age of Artificial Intelligence: A Comparative Analysis of US, EU, and UK Regulatory Frameworks," *AI and Ethics* 2 (August 2022), pp. 377–87.
[4] "Watching the Border," *Economist*, February 19, 2022, pp. 19–21.
[5] By April of 2022, Clearview AI had created more than two hundred accounts and conducted more than five thousand searches. Clearview AI also translated its app into Ukrainian. See Kashmir Hill, "Facial Recognition Goes to War," *New York Times*, April 7, 2022, www.nytimes.com/2022/04/07/technology/facial-recognition-ukraine-clearview.html.
[6] On the justification for social media intelligence in liberal democracies, see Kira Vrist Rønn and Sille Obelitz Søe, "Is Social Media Intelligence Private? Privacy in Public and the Nature of Social Media Intelligence," *Intelligence and National Security* 34, no. 3 (2019), pp. 362–78.
[7] Fabre, *Spying through a Glass Darkly*, p. 3.
[8] Self-defensive harm is of course subject to constraints such as necessity, proportionality, and success, and there might be circumstances in which harm to an individual may result as a consequence of justified collateral damage. See Fabre, *Spying through a Glass Darkly*, p. 25.
[9] Valerie Hopkins, "A Nation of Spy-Catchers: Fear of Saboteurs Has Ukrainians on Edge," *New York Times*, March 31, 2022, www.nytimes.com/2022/03/31/world/europe/ukraine-spies-saboteurs.html.
[10] Alexander J. Motyl, "From Ukraine, with Loot," *Hill*, April 26, 2022, thehill.com/opinion/international/3460498-from-ukraine-with-loot.
[11] Here I am highlighting the deterrent effects of punishments to illuminate another dimension of threat prevention, but sheer retribution is another justification for punishment, which FRT can also help attain.
[12] Brett Forrest, "Ukraine Has Deployed 1,000 People to Investigate Alleged War Crimes in Bucha," *Wall Street Journal*, updated April 12, 2022; and Jeff Wise, "The Hunt for the Butchers of Bucha," *New York Magazine*, April 8, 2022.
[13] Mykhailo Fedorov, quoted in Johana Bhuiyan, "Ukraine Uses Recognition Software to Identify Russian Soldiers Killed in Combat," *Guardian*, March 24, 2022.
[14] See Art. 34(1), "Remains of Deceased," in Diplomatic Conference on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts, "Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of Non-International Armed Conflicts (Protocol I)," June 8, 1977.
[15] With respect to enemy soldiers, this duty may be interpreted as a negative duty not to intervene in the identification of fallen soldiers. With respect to the belligerent party's own soldiers/citizens, it would be more appropriate to describe it as a positive duty to conduct such identifications.
[16] Pauline Boss, *Ambiguous Loss: Learning to Live with Unresolved Grief* (Cambridge, Mass.: Harvard University Press, 1999).
[17] Martha C. Nussbaum, *Creating Capabilities: The Human Development Approach* (Cambridge, Mass.: Belknap, 2013).
[18] Marcus Smith and Seumas Miller, "The Ethical Application of Biometric Facial Recognition Technology," *AI & Society* 37 (March 2022), pp. 167–75.
[19] Paresh Dave and Jeffrey Dastin, "Ukraine Has Started Using Clearview AI's Facial Recognition during War," Reuters, March 14, 2022.
[20] Fabre, *Spying through a Glass Darkly*, p. 216.
[21] Avia Pasternak, "Voluntary Benefits from Wrongdoing," in "Benefiting from Injustice," special issue, *Journal of Applied Philosophy* 31, no. 4 (November 2014), pp. 377–91.
[22] Avia Pasternak and Zofia Stemplowska, "Are Severe Sanctions on Russia Morally Justified?," *New Statesman*, April 19, 2022.

[23] Selinger and Leong, "The Ethics of Facial Recognition Technology."
[24] Frederick Schauer, "Slippery Slopes," *Harvard Law Review* 99, no. 2 (December 1985), p. 361.
[25] Scott Robbins, "Facial Recognition for Counter-Terrorism: Neither a Ban nor a Free-for-All," in Adam Henschke, Alastair Reed, Scott Robbins, and Seumas Miller (eds.), *Counter-Terrorism, Ethics and Technology: Emerging Challenges at the Frontiers of Counter-Terrorism* (Cham, Switzerland: Springer, 2021), pp. 89–104. See also Carissa Véliz, *Privacy Is Power: Why and How You Should Take Back Control of Your Data* (London: Random House, 2020).
[26] Darian Meacham and Martin Gak, "Does Facial Recognition Tech in Ukraine's War Bring Killer Robots Nearer?," openDemocracy, March 30, 2022.
[27] Brett Frischmann and Evan Selinger, *Re-Engineering Humanity* (New York: Cambridge University Press, 2018), p. 39.
[28] Bhuiyan, "Ukraine Uses Recognition Software to Identify Russian Soldiers Killed in Combat."
[29] Evan Greer, quoted in Hill, "Facial Recognition Goes to War."
[30] Thanks to Cécile Fabre for suggesting this particular idea.
[31] James Clayton, "How Facial Recognition Is Identifying the Dead in Ukraine," BBC News, April 13, 2022.
[32] Michael Skerker, "Moral Concerns with Cyberespionage: Automated Keyword Searches and Data Mining," in Fritz Allhoff, Adam Henschke, and Bradley Jay Strawser (eds.), *Binary Bullets: The Ethics of Cyberwarfare* (Oxford: Oxford University Press, 2016). That the risks are lower does not mean, obviously, that they can be ruled out.
[33] According to the World Justice Project's "Rule of Law Index" for 2021, prewar Ukraine had low state capabilities in the "Regulatory Enforcement," "Government Powers," and "Absence of Corruption" categories. On a scale from 0 to 1, it scored 0.44, 0.47, and 0.33, respectively: significantly lower than nations like the U.S. or the U.K. See the "Overall Index Score" table on the World Justice Project website at worldjusticeproject.org/rule-of-law-index/global/2021/table.
[34] At least three months into the war, Clearview AI had not revoked any access to the service. See Dina Temple-Raston and Sean Powers, "At War with Facial Recognition: Clearview AI in Ukraine," *The Record*, May 16, 2022, therecord.media/at-war-with-facial-recognition-clearview-ai-in-ukraine.
[35] Robbins, "Facial Recognition for Counter-Terrorism."
[36] Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (New York: PublicAffairs, 2019).
[37] On the concerns about digital governance and the legitimacy deficit, see Zuboff, *The Age of Surveillance Capitalism*; Véliz, *Privacy Is Power*; and Claire Benn and Seth Lazar, "What's Wrong with Automated Influence," *Canadian Journal of Philosophy* 52, no. 1 (January 2022), pp. 125–48.
[38] David C. Cornett, David S. Bolme, Dawnie W. Steadman, Kelly A. Sauerwein, and Tiffany B. Saul, "Effects of Postmortem Decomposition on Face Recognition" (conference report, 2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems [BTAS], Tampa, Florida, 2019), pp. 1–8, ieeexplore.ieee.org/document/9185971/citations#citations.
[39] International Committee of the Red Cross, Art. 16, "Wounded and Sick: I. General Protection," Convention 4, "The Geneva Conventions of August 12, 1949."
[40] Art. 34(1), Diplomatic Conference on the Reaffirmation and Development of International Humanitarian Law, "Protocol Additional to the Geneva Conventions of 12 August 1949."
[41] Sarah Ashbridge, "Digital Dignity in Death: Are the Geneva Conventions Fit for Purpose in the Age of Social Media?," Royal United Services Institute, March 29, 2022, rusi.org/explore-our-research/publications/commentary/digital-dignity-death-are-geneva-conventions-fit-purpose-age-social-media. Ukrainian authorities have posted on social media platforms a steady stream of extremely graphic images illustrating the atrocities of war, inviting Russians to identify a missing loved one. See Drew Harwell, "The Gory Online Campaign Ukraine Hopes Will Sow Anti-Putin Dissent Probably Violates the Geneva Conventions," *Washington Post*, March 3, 2022.
[42] Harwell, "The Gory Online Campaign Ukraine Hopes Will Sow Anti-Putin Dissent Probably Violates the Geneva Conventions."
[43] At least in a plausible reading of the Conventions, since for obvious reasons they are silent on issues pertaining to social media or FRT.
[44] On the principle of discrimination in just intelligence theory, see Ross W. Bellaby, *The Ethics of Intelligence: A New Framework* (London: Routledge, 2014), pp. 35–37; and Kevin Macnish, "An Eye for an Eye: Proportionality and Surveillance," *Ethical Theory and Moral Practice* 18, no. 3 (June 2015), pp. 529–48. In the just war tradition, the classic references are Michael Walzer, *Just and Unjust Wars: A Moral Argument with Historical Illustrations* (New York: Basic Books, 2006) and Jeff McMahan, *Killing in War* (Oxford: Oxford University Press, 2009).

[45] Seth Lazar, "War," in *Stanford Encyclopedia of Philosophy* archive, May 3, 2016, plato.stanford.edu/archives/spr2020/entries/war/.

Abstract: The use of facial recognition technology (FRT) as a form of intelligence has recently made a prominent public appearance in the theater of war. During the early months of Russia's invasion of Ukraine, Ukrainian authorities relied on FRT as part of the country's defensive activities, harnessing the technology for a variety of purposes, such as unveiling covert Russian agents operating amid the Ukrainian population; revealing the identity of Russian soldiers who committed war crimes; and even identifying dead Russian soldiers. This constellation of uses of FRT—in a war increasingly waged on the digital and information front—warrants ethical examination. The essay discusses some of the most serious concerns with FRT in the context of war, including the infringement of informational privacy; the indiscriminate and disproportionate harms it may inflict, particularly when the technology is coupled with social media intelligence; and the potential abuse of the technology once the fog of war dissipates. Some of these concerns parallel those to be found in nations that are not engulfed in war, but others are unique to war-torn settings.

Keywords: facial recognition, mass surveillance, espionage, war, social media, Russia, Ukraine

https://doi.org/10.1017/S0892679423000151 Published online by Cambridge University Press