Essay 1 Outline

Topic

Copyright Issues in Training an Image Recognition AI for Student Projects

Dilemma

Moral question: When building image-recognition AI for student projects, scraping copyrighted images from the internet could significantly improve model performance. So, is it morally permissible to use those images for training without permission?

- **Only focus on performance**: Scraping large quantities of high-quality online images, many of which appear to be copyrighted
    - **Benefits**: Higher accuracy, stronger project results, and more learning and demonstration value.
    - **Costs**: Usually, creators' work was used without consent and creative labour was treated as “free raw material.” The project’s ownership was blurred.
- Respecting the owner's right: Only use self-collected images, school datasets, or images with open licenses.
    - **Benefits**: Clear permission, lower ethical risk.
    - **Costs**: Smaller datasets lead to lower-quality and weaker model performance, which might reduce educational value.

Real-world case

TechCrunch reported that some YouTube creators filed a lawsuit against Snap. They allege Snap trained key AI models on their original content without explicit permission, which could amount to copyright infringement and misuse of their intellectual property.

Thesis

In general, students should not train an image recognition model on copyrighted images without explicit permission. Although it may be unrealistic to avoid all potentially copyrighted content, students should choose reasonable, feasible alternatives when available; if no acceptable alternative exists, they should narrow the project scope rather than cross the ethical boundary.

Argument for your thesis

1. If performance gains in training depend on others' work, then using that work without permission is an unfair appropriation of their labour.
2. The value of a student project lies not only in its performance but also in the learning process and in the ability to explain and justify its results. Using data from an unclear source or without permission undermines the project’s suitability for public sharing.
3. If non-consensual scraping and training become normal, creators may respond with stronger defensive and adversarial measures, which would increase lawsuits and conflict

and reduce public trust in AI. After a long time, these broader costs may outweigh the short-term performance benefits of a student project.

Consider what a relativist would recommend

A normative moral relativist might say that whether scraping is “right” depends on the norms of the relevant group. For example:

- In some technical communities, people treat all public content as acceptable training data, so this community may be judged to be scraping permissible.
- In some communities that emphasise creators’ rights, this may be judged wrongful.

This relativist position is appealing because norms and practices vary widely across cultures and communities, making it seem as though there is no single universal answer.

Reject relativism for objectivism

A common line of reasoning used by relativism can be stated as:

- **Empirical premise:** Different groups have different moral norms.
- **Normative conclusion**: Therefore, there is no objective moral truth.

This argument is false because it directly jumps from an empirical fact to a normative conclusion. Disagreement can not show that there is no fact of the matter. To reach a normative conclusion, we must supply additional value premises and argumentation, rather than treating disagreement itself as decisive. Therefore, I reject normative moral relativism and defend moral objectivism.

If relativism were true, it would have at least two serious consequences:

1. We could not genuinely criticise other groups’ behaviour, as moral judgment could simply become checking local customs.
2. This will make moral progress harder to explain, because "better" simply means "currently popular" under relativism.

In this case, if “whatever a group accepts is right” were true, then infringements of creators’ works should be downgraded to mere conflicts of preference between different groups, rather than regarded as issues of legitimacy open to serious evaluation. However, the TechCrunch report shows that these controversies cannot be dismissed by a simple “each group has its own standards.”

Consider what a Utilitarian would advise

Utilitarianism would count the benefits for all people affected, not only the student team:

- **Short-term benefits**: The student project has better model performance, learning experience, and demonstrations.
- **Potential harms**: Creators' interests could be damaged, public trust might be reduced, and conflict could increase.

In a typical student project, the short-term benefits are mainly limited to a small group, while the long-term harm affects a much larger group. Therefore, a utilitarian would probably recommend a general rule favouring permission-first datasets, which could include open-licensed or

school-provided sources, and, when data appears insufficient, the approach might suggest adjusting project scope and goals to maximise overall welfare over time.

Objection to your argument + Reply

Objection

Students' resources are limited, and it is hard for them to have enough licensed images. Also, model training is internal and does not directly republish the images, so it does not really harm creators.

Reply

- No ought without can: if something is genuinely impossible, it cannot be required. Feasibility does not mean morality can be crossed.
- A more reasonable response is to narrow the task, rely on permission data, rather than externalising costs onto creators by using copyrighted works without permission.

Reference

[YouTubers sue Snap for alleged copyright infringement in training its AI models]

AI acknowledgment

Because I had never created an outline like this before:

1. First, I wrote a long paragraph for ChatGPT describing what I wanted to say (basically, what I wrote in the lab and my understanding of the essay requirements).
2. Then it reorganised the content into the current sections and provided many revision suggestions to better align with the essay requirements.
3. Based on that, I made further edits because sometimes ChatGPT shifted my ideas too far from what I originally meant, and some parts no longer reflected what I wanted to express. So I spent time revising those sections.

I hope these changes make my essay more logical and better structured.
In addition, ChatGPT was used to translate the statement above to improve grammatical accuracy.