

6. WYNIKI

Wstęp

Ten rozdział służy prezentacji wyników pracy zespołu. Najważniejszym elementem projektu było wytworzenie modelu generującego mapy. Wszystkie inne elementy projektu (szczególnie te związane z pozyskiwaniem i przetwarzaniem danych) ostatecznie prowadziły do wyszkolenia generatora i mają w nim swoje odbicie. W związku z tym ten rozdział jest poświęcony w całości procesowi treningu, zachowaniu utworzonych modeli oraz właściwościom wytwarzanych przez nie obrazów.

Zostały tutaj również przedstawione rezultaty różnych eksperymentów przeprowadzonych w ramach treningu modeli. Szczególnie istotną zmianą okazała się być modyfikacja rozmiarów filtrów konwolucyjnych w generatorze, w związku z czym właściwości generowanych obrazów zostały przedstawione w jej kontekście.

6.1. Szybkość treningu

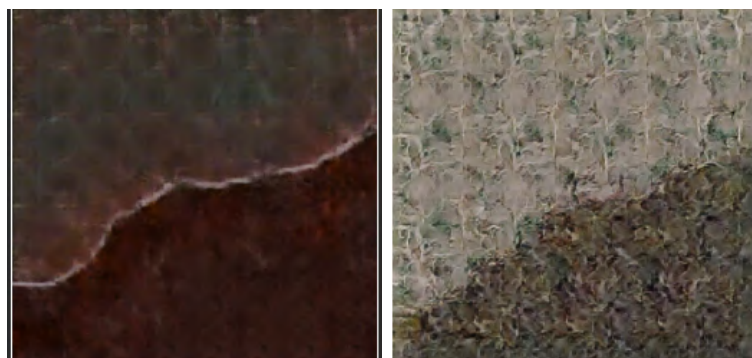
Trening modeli prowadzony był na maszynie wyposażonej w procesor Ryzen 5 3600, 16 GB pamięci RAM oraz kartę graficzną Nvidia RTX 4060 z 8 GB pamięci VRAM. Zainstalowano sterowniki karty graficznej w wersji 550.120.

Trening prowadzony był na pojedynczych obrazach - każda partia (batch) zawierała tylko jeden obraz. Średni czas treningu z domyślnymi parametrami (czyli takimi, jak opisano w 5.4) na pojedynczym obrazie wynosił 0,069 sekundy. Czas ten wyznaczono poprzez uśrednienie czasu treningu na pierwszych 10000 obrazach podczas losowo wybranej sesji treningu. Cały trening modelu był za każdym razem prowadzony przez 40000 kroków, co zajmowało przeciętnie około 45 minut.

We wszystkich przeprowadzonych eksperymentach związanych ze zmianą parametrów modeli czas treningu nie zmieniał się w sposób zauważalny (czas za każdym razem mierzony był tak, jak opisano wyżej). Wyjątek stanowiły próby, podczas których zmieniane były rozmiary filtrów konwolucyjnych generatora. Był to spodziewany rezultat - powiększanie filtrów konwolucyjnych znacznie zwiększa liczbę parametrów modelu. Podczas próby z filtrami o wymiarach 6x6 czas treningu na pojedynczym obrazie wyniósł 0.131 sekundy. W związku z tym czas pełnego treningu wydłużył się do około 90 minut (liczba przykładów pozostawała niezmienną).

Podczas próby z filtrami o rozmiarze 8x8 średni czas treningu na pojedynczym obrazie wzrósł już do 0,206 sekundy, co poskutkowało wydłużeniem treningu do ponad dwóch godzin.

Próba, w której filtry miały wymiary 2x2 trwała około 25 minut, przy średnim czasie na obraz równym 0,04 sekundy.



rys. 6.1. Porównanie obrazów wytworzonych przez model wytrenowany na mapach podziemi (po lewej) oraz przez model wytrenowany na mapach powierzchni (po prawej). Dla obu map obrazy wejściowe były niemal identyczne.

6.2. Jakość wyników a zbiór danych

6.2.1. Podział zbioru danych

Divinity: Original Sin 2 to gra o bardzo zróżnicowanych lokacjach, które utrzymane są w charakterystycznych stylach. Co za tym idzie, wygląd danego typu terenu może różnić się z lokacji na lokację.

W początkowych fazach rozwoju projektu do treningu wykorzystywany był cały dostępny zbiór danych. Rezultaty nie były jednak zadowalające: pewne typy powierzchni (w szczególności skała), przybierały na obrazach wyjściowych nijaki, mało interesujący wygląd, stanowiący uśrednienie wszystkich stylów występujących wśród danych treningowych.

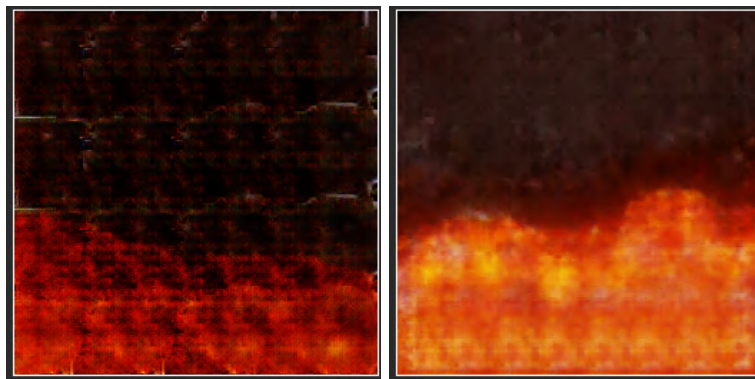
Aby rozwiązać ten problem, należałoby wprowadzić dodatkowe etykiety typów terenu, zależne od stylu (na przykład "skała-podziemia" i "skała-góry", w miejsce pojedynczej etykiety "skała"). Alternatywą było podzielenie zbioru danych na podzbiory względem stylu i wytrenowanie kilku modeli, spośród których każdy tworzyłby obrazy w innym stylu. Ostatecznie wybrano tę drugą opcję. Do uzyskania oczekiwanych rezultatów wystarczył podział na dwa zbiory danych: "Podziemia" oraz "Powierzchnia".

Na rysunku (6.1) ilustracjach przedstawiono różnice w obrazach generowanych przez dwa modele. Pierwszy obraz został stworzony przez model wytrenowany na zbiorze map podziemi, zaś drugi stanowi dzieło modelu trenowanego na zbiorze map powierzchni. Oba modele otrzymały niemal identyczny obraz wejściowy, który w lewym górnym rogu miał powierzchnię kamienną stworzoną przez człowieka, zaś w prawym dolnym rogu skałę.

6.2.2. Selekcja danych

Mimo że dane zostały rozdzielone względem stylu, rezultaty podczas treningu na zbiorze map podziemi wciąż pozostawały niezadowalające. Jest to bezpośrednio związane z problemami opisanymi w sekcji "Problematiczne typy terenu".

W celu poprawy jakości wyników, podjęto wiele prób wybiórczego treningu, a podczas każdej z nich wykluczano część dostępnych danych. Przykładowo, przeprowadzony został trening wyłącznie



rys. 6.2. Zestawienie obrazów wytworzonych przez model wytrenowany na zbiorze zawierającym obrazy niedoświetlone (powyżej) oraz przez model trenowany na przefiltrowanych danych (poniżej)

na podziemiach niezawierających elementów wytworzonych przez człowieka. Takie próby jednak również dawały słabe rezultaty, przypuszczalnie ze względu na zbyt małą różnorodność danych. Wyniki uległy zdecydowanej poprawie dopiero, kiedy ze zbioru danych wykluczony został podzbiór obrazów ciemnych, które sprawiały wrażenie niedoświetlonych (rysunek 6.2).

Mimo to, jakość wyników produkowanych przez model dedykowany podziemiom jest zauważalnie niższa, niż w przypadku modelu tworzącego mapy powierzchni. Przejawia się to mniejszą liczbą terenów, które udało się odwzorować w przekonujący sposób oraz ich niższą szczegółowością. Może to być spowodowane faktem, iż zbiór ten zawierał mniej danych, ze względu na niewielką liczbę odpowiednio dużych map podziemi dostatecznej jakości, które dało się pozyskać z Divinity: Original Sin II.

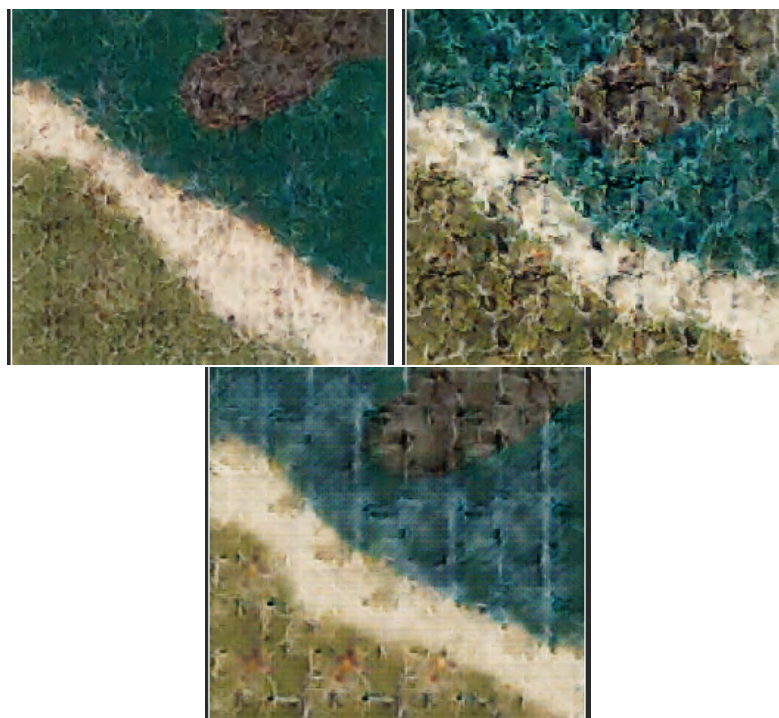
6.3. Wyniki eksperymentów

6.3.1. Zmiana wag funkcji straty

Jak napisano w 5.4.5, podczas treningu wykorzystywane są dwa rodzaje straty - wynik pracy dyskriminatora oraz strata bezwzględna, która stanowi różnicę między wartościami pikseli obrazu wygenerowanego a docelowego (to znaczy tym, na podstawie którego stworzono obraz wejściowy).

Obie wartości funkcji straty obliczane są w ramach każdego kroku uczenia sieci, jednak ich wagi nie są identyczne. Autorzy architektury pix2pix w swoim artykule zalecają, aby stosunek wag tych wynosił 100:1, z przewagą dla bezwzględnej różnicy między obrazami. To oznacza, że ocena dyskriminatora jest stukrotnie mniej ważna od różnicy między obrazem wygenerowanym a docelowym.

Podjęty został zatem szereg prób, mających zweryfikować, czy w przypadku problemu generowania map zmiana tego stosunku nie przyniosłaby bardziej interesujących i mniej powtarzalnych rezultatów. Jak się okazało, jakość wyników spada wraz ze spadkiem znaczenia funkcji straty bezwzględnej (rysunek 6.3). Generowane obrazy stają się mniej szczegółowe i bardziej rozmyte. Częściej generowane są na nich jasne linie, które na oryginalnych mapach stanowią granice obszaru dostępnego dla gracza. Przy szczególnie niskiej wadze przypisanej funkcji straty bezwzględnej, obraz wyjściowy zostaje przez te jasne linie podzielony na równej wielkości fragmenty o nieregularnych (ale identycznych)



rys. 6.3. Przykładowe wyniki treningu w zależności od stosunku wag obu składowych funkcji straty. U góry po lewej: 1:100, z przewagą różnicy bezwzględnej (parametry domyślne.). U góry po prawej: 9:1, z przewagą różnicy bezwzględnej. Na dole: 1:1.

kształtach, które pozostają niezmiennione, niezależnie od podanego obrazu wejściowego.

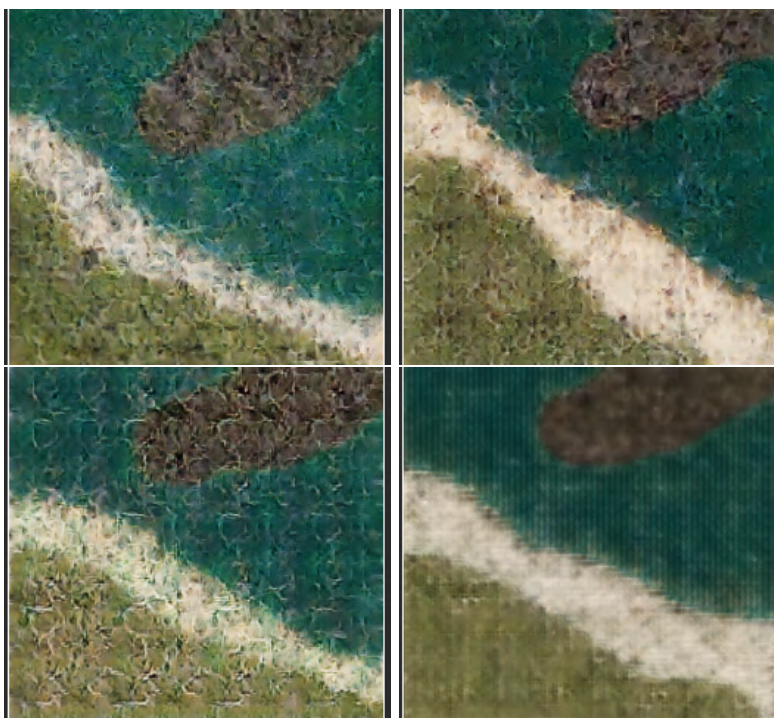
6.3.2. Zmiana okna dyskryminatora

Rozmiar okna dyskryminatora to kolejny parametr, którego wpływ na jakość generowanych obrazów był obiektem eksperymentów. Jak opisano w 5.4.2, dyskryminator nie ocenia podawanych mu obrazów w całości. Zamiast tego dzieli je na mniejsze, kwadratowe fragmenty i ocenia jakość każdego z osobna, a ocenę danego obrazu stanowi średnia ocena wszystkich składających się na niego fragmentów.

Rozmiar fragmentów ocenianych indywidualnie przez dyskryminator można było łatwo modyfikować. Domyślnie (zgodnie z zaleceniami autorów publikacji) miały one wymiary 70 na 70 pikseli. Przeprowadzonych zostało kilka prób, podczas których sprawdzono szereg możliwych wielkości okna. Rozmiar okna był jedynym parametrem modyfikowanym podczas tych prób. Sieć za każdym razem była trenowana na takiej samej liczbie przykładów (40000). Wyniki przedstawia rysunek (6.4)

Zmniejszenie okna dyskryminatora nie przyniosło znaczących zmian w jakości tworzonych obrazów - uzyskiwane obrazy są an podobnym poziomie szczegółowości. Występują między nimi drobne różnice, ale są one natury estetycznej.

Zwiększenie okna do rozmiaru 128x128 poskutkowało natomiast wyraźnym obniżeniem jakości obrazów. Na całej powierzchni każdego utworzonego obrazu pojawiają się na nich białe, zakrzywione linie i okręgi. Przypominają one te linie, które na oryginalnych mapach wytyczają granice obszarów dostępnych dla graczy.



rys. 6.4. Przykładowe wyniki treningu w zależności od rozmiaru okna dyskryminatora. Lewy górny róg: 32x32. Prawy górny róg: 70x70 (rozmiar domyślny). Lewy dolny róg: 128x128. Prawy dolny róg: 256x256 (w tym przypadku dyskryminator ocenia cały obraz).

Test, w ramach którego okno miało wymiary 256x256 (czyli oceniany był cały obraz), przyniosła największy spadek jakości. Zachowane zostały kolory poszczególnych powierzchni, jednak wszystkie tekstury uległy rozmyciu, co uczyniło je całkowicie nieczytelnymi.

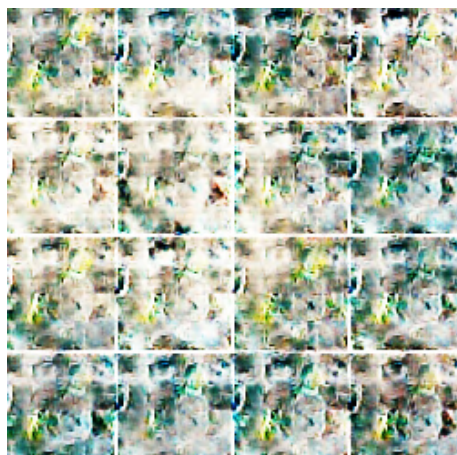
6.3.3. Modyfikacja wielkości filtrów konwolucyjnych

Zmiana wymiarów filtrów konwolucyjnych okazała się być najbardziej znaczącą modyfikacją, która została wprowadzona do modelu generatora. Była to jednocześnie jedyna modyfikacja oryginalnej architektury pix2pix, która pozwoliła realnie poprawić jakość uzyskiwanych wyników.

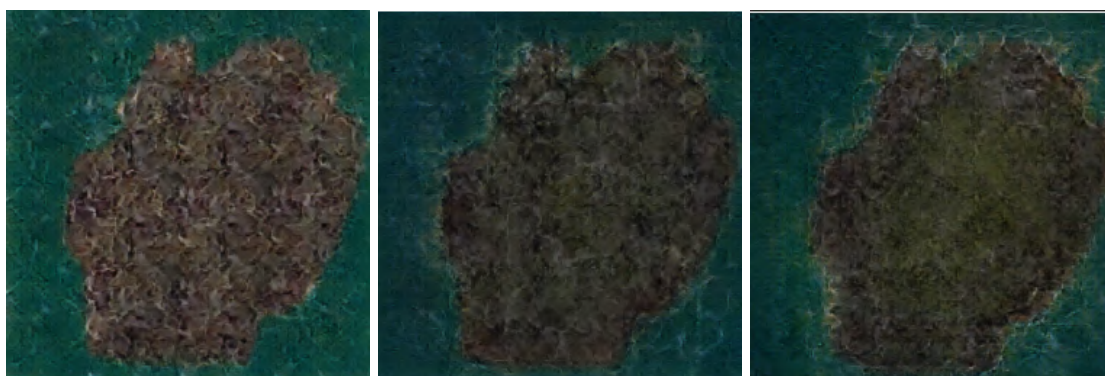
Najmniej imponujące wyniki miała próba, w której rozmiar filtrów został zredukowany do 2x2. W tym przypadku odnotowany został drastyczny spadek jakości generowanych obrazów - stały się one praktycznie nieczytelne. Niezależnie od wprowadzonego obrazu wejściowego, cały obraz jest pokryty siatką białych, kwadratowych ramek. Dodanie terenu na obrazie wejściowym skutkuje jedynie delikatną zmianą koloru w tle, za białymi artefaktami (rysunek 6.5).

Próby podczas których filtry konwolucyjne były większe (6x6 i 8x8) przyniosły poprawę jakości tworzonych map. Zaobserwowane różnice były podobne dla obu przypadków, jednak dla filtrów o wymiarach 8x8 ich natężenie było dużo większe. Różnice między tymi modelami widać na rysunku 6.6.

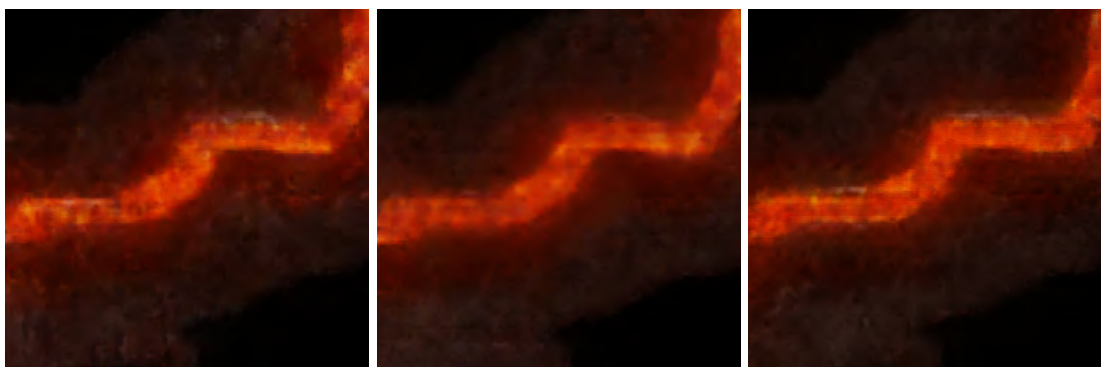
Podstawowa różnica polegała na większej szczegółowości generowanych tekstur terenów. Model zaczął także lepiej reagować na kontekst, w którym pojawiają się poszczególne typy terenów. Tereny zaczęły również mieszać się na krawędziach w bardziej naturalny sposób.



rys. 6.5. Wynik pracy generatora o filtrach rozmiaru 2x2. Takie ograniczenie złożoności modelu poskutkowało drastycznym spadkiem jakości.



rys. 6.6. Porównanie wyników modeli o trzech różnych stopniach złożoności, uruchomionych na identycznym obrazie wejściowym. Od lewej: wyniki pracy modeli o filtrach rozmiaru odpowiednio 4x4, 6x6 oraz 8x8



rys. 6.7. Porównanie wyników modeli generujących mapy podziemi o trzech różnych stopniach złożoności, uruchomionych na identycznym obrazie wejściowym. Od lewej: wyniki pracy modeli o filtrach rozmiaru odpowiednio 4x4, 6x6 oraz 8x8

Powyższe usprawnienia dotyczą jednak tylko modelu produkującego mapy powierzchni. Model służący do tworzenia map podziemi nie skorzystał w znaczący sposób na powiększeniu filtrów konwolucyjnych. Jest to przypuszczalnie spowodowane niedostateczną ilością dostępnych danych. Widoczne są niewielkie różnice (rysunek 6.7), jednak ograniczają się one do wygładzenia tekstur pokrywających poszczególne typy terenów, co pozbawiło je szczegółów i w praktyce obniżyło ich walory estetyczne. W związku z tym, w dalszej części tego rozdziału, wszelkie opisy zachowania bardziej skomplikowanej wersji modelu dotyczyć będą jedynie modelu tworzącego mapy powierzchni.

Dokładne różnice w wynikach produkowanych przez poszczególne wersje modelu omówione zostały dokładnie w kolejnych sekcjach.

6.4. Generowane Obrazy

6.4.1. Tekstury

Wprowadzone na wejściu plamy kolorów zostają pokryte teksturami reprezentującymi odpowiednie tereny (rysunek 6.8). Mimo że tereny nie są pokryte powtarzającymi się obrazami, na ich powierzchniach można zaobserwować regularne, powtarzalne wzory. Obniża to wizualną jakość map i przypomina nieco problem tilingu tekstur, pojawiający się przy tworzeniu gier komputerowych.

Powtarzające się wzory bywają bardziej lub mniej widoczne w zależności od rodzaju terenu. Problem powtarzalnych wzorów jest dużo mniej widoczny w przypadku modeli o większych filtrach. Tutaj tereny są mniej jednolite, a na ich powierzchniach można zauważyć elementy które nieco je urozmaicają - przykładowo na powierzchniach skalnych pojawia się mech, który jednak nie zajmuje całej ich powierzchni. Jeśli jednak utworzymy obraz zawierający dużą połąć terenu jednego typu, to wciąż zauważymy pewne powtarzające się wzory, będą one jednak dużo drobniejsze (rysunek 6.9).

6.4.2. Problematyczne typy terenu

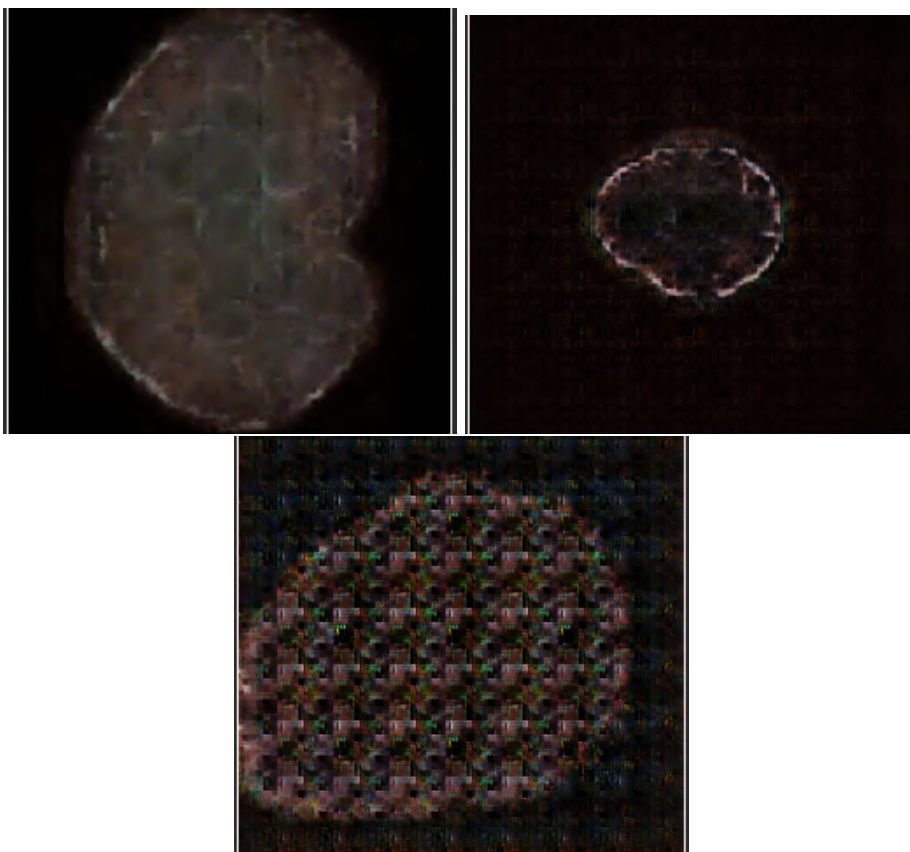
Podczas procesu uczenia, generator uczył się odwzorowywać pewne typy terenów dużo skuteczniej niż inne. Problemy pojawiały się przede wszystkim w przypadku terenów, dla których dane treningowe były bardziej zróżnicowane: było tak chociażby w przypadku powierzchni tworzonych przez



rys. 6.8. Przykładowe powierzchnie tworzone przez generator. Kolejno, od lewej i od góry: kamień wytworzony przez człowieka na mapie powierzchni, skała na mapie powierzchni, lawa na mapie podziemi, trawa na mapie powierzchni i kamień wytworzony przez człowieka na mapie podziemi.



rys. 6.9. Demonstracja powierzchni generowanych przez najbardziej złożoną wersję modelu (8x8). Od lewej: trawa, skała oraz kamień pochodzenia ludzkiego.

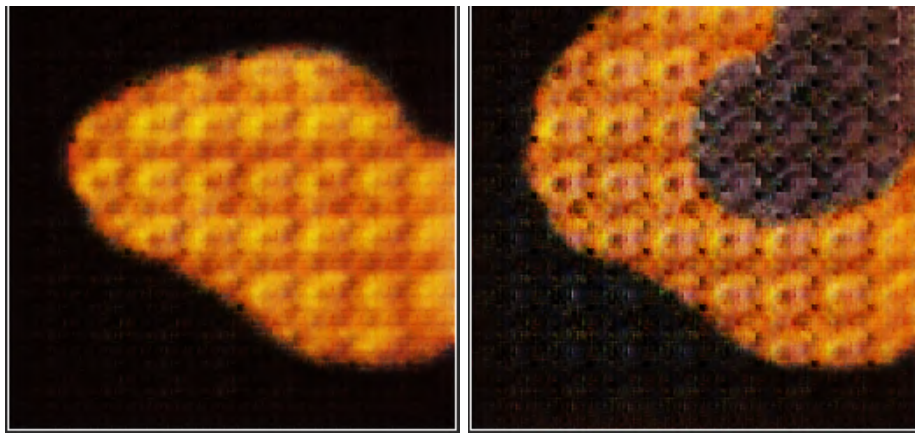


rys. 6.10. Przykłady problematycznego terenu - kamienia stworzonego przez człowieka na mapach podziemi. U góry po lewej: teren względnie poprawny, jednak niewyraźny i z widocznymi białymi śladami. U góry po prawej oraz na dole: obrazy pochodzące z treningu na nieprzefiltrowanym zbiorze danych - jest to bardziej pesymistyczny spośród dwóch opisanych scenariuszy

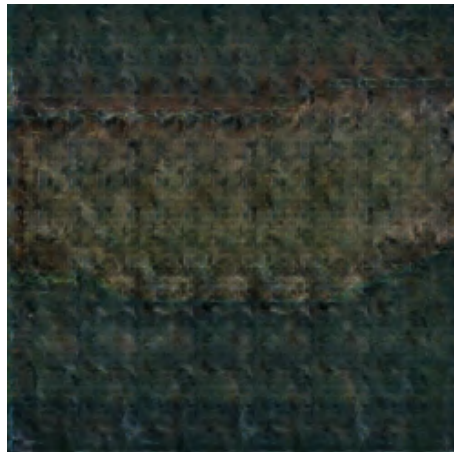
człowieka, takich jak bruk czy posadzki (rysunek 6.10).

Takie tereny w najlepszym wypadku były odtwarzane jako rozmyte, szare lub brunatne plamy, pokryte ledwo czytelnym geometrycznym motywem. W niektórych próbach odtworzone zostały jasne linie, które na mapach gry wyznaczają granice terenów dostępnych dla postaci gracza.

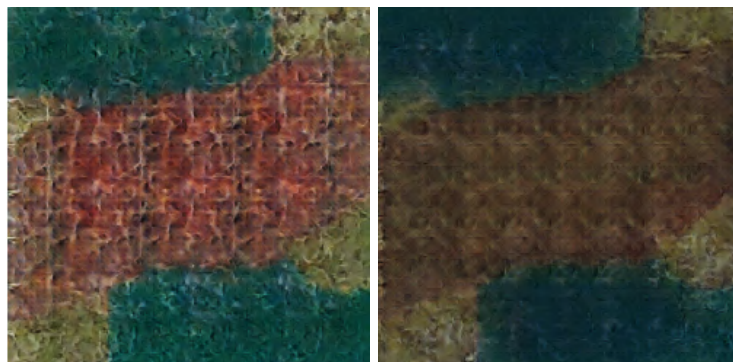
W niektórych zbiorach danych, wspomniane wyżej tereny były bardziej problematyczne i na wyjściu przybierały postać ciemnych plam, losowo nakrapianych jasnymi, jaskrawymi pikselami. Naniesienie tych typów powierzchni na obraz wyjściowy skutkowało obniżeniem jakości całego obrazu wyjściowego (nawet w miejscach pozbawionych trudnego terenu). Obniżenie jakości polegało na pojawieniu się jasnego ziarna na wszystkich typach terenu oraz na obniżeniu czytelności tekstur (rysunek 6.11). Istniał również typów terenów, których model w ogóle nie nauczył się odwzorowywać (rysunek 6.12), bądź nauczył się je odwzorowywać w sposób niezadowalający. Do takich terenów należą między innymi powierzchnie drewniane, czy ścieżki tworzone przez człowieka. Tym, co łączy te tereny, jest fakt, iż w zbiorze danych znajdowało się względnie mało przykładów takich terenów, a ponadto były one bardzo zróżnicowane. Bardziej skomplikowane modele poradziły sobie lepiej. Udało im się odwzorować trudne tereny (w szczególności te antropogeniczne, rysunek 6.13) w bardziej przekonujący sposób. Mimo to, istniały pewne tereny, których nawet im nie udało się odtworzyć.



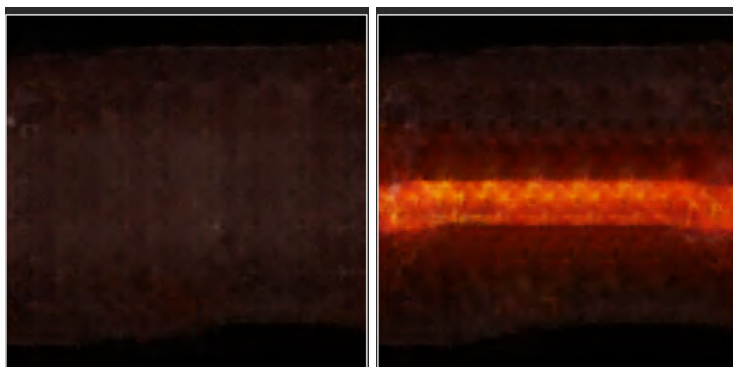
rys. 6.11. Przykład tego, w jaki dodanie plamy problematycznego terenu psuje jakość wszystkich innych terenów przedstawionych na mapie. Po lewej: mapa bez trudnego terenu. Z prawej: wynik dostarczony przez ten sam model po dodaniu problematycznego terenu



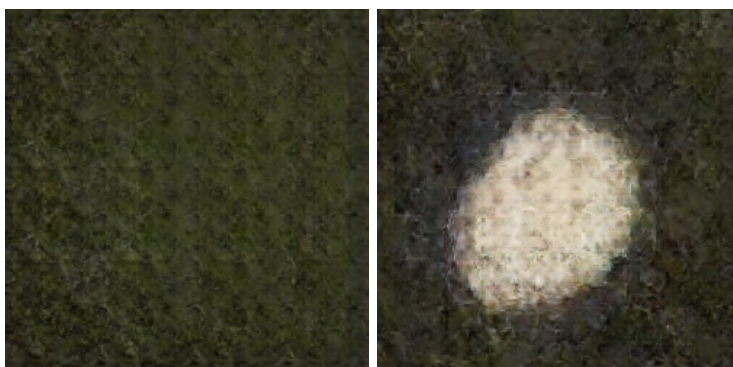
rys. 6.12. Przykład terenów, których nie udało się nauczyć żadnemu z modeli. Ta mapa ma przedstawiać ścieżkę stworzoną przez człowieka na tle gleby. Podobieństwo jest zerowe.



rys. 6.13. Przykład problematycznego terenu: drewniany most. Po lewej dzieło podstawowego generatora. Po prawej ten sam obraz wejściowy ale przetworzony przez generator o filtrach wielkości 6x6.



rys. 6.14. Porównanie tej samej mapy przed i po dodaniu rzeki lawy. Na drugiej ilustracji wyraźnie widać pomarańczową poświatę, padającą na skałę



rys. 6.15. W bardziej złożonych modelach powierzchnie potrafią przybierać zupełnie inny wygląd, w zależności od kontekstu w którym się pojawiają. Tutaj: mech znika z kamienia, kiedy tylko w okolicy pojawi się piasek.

6.4.3. Interakcja między typami terenu

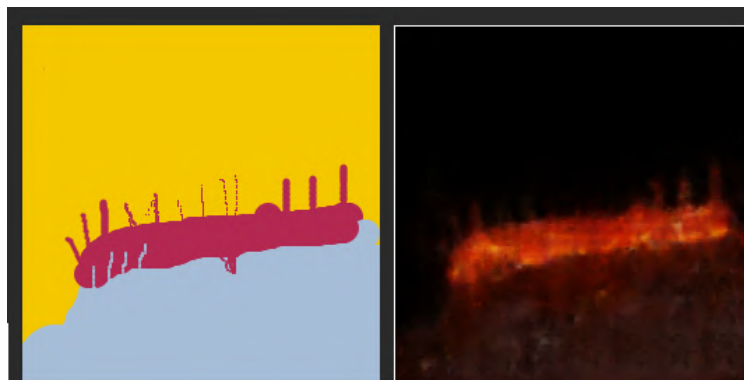
Na obrazach wynikowych można zauważyć drobne szczegóły, świadczące o tym, że tereny nie są rysowane zupełnie niezależnie od siebie, a model bierze pod uwagę występujące między nimi relacje przestrzenne. Dzięki temu teren może posiadać delikatnie zmieniony wygląd, w zależności od innych sąsiadujących z nim terenów.

Najbardziej widoczne jest to w przypadku lawy (rysunek 6.14), która rzuca poświatę na otaczające ją tereny, sprawiając, że ich krawędzie sąsiadujące z lawą są jaśniejsze i zabarwione na pomarańczowo. Ta właściwość jest znacznie lepiej widoczna w przypadku bardziej złożonych modeli o większych filtrach. Biorą one pod uwagę sąsiedztwo między poszczególnymi typami terenu w znacznie większym stopniu, niż modele prostsze. Jest to spodziewany wynik, jako że głównym zadaniem jest właśnie wychwytywanie zależności przestrzennych między pikselami.

W ten sposób, na przykład, skały domyślnie porośnięte są mchem albo rzadką roślinnością. Kiedy jednak w okolicy powierzchni skalnej znajdzie się piasek, mech znika. Ta konkretna właściwość jest prawdopodobnie spowodowana faktem, że skałom na oryginalnych mapach z gry zazwyczaj towarzyszy roślinność, jednak te znajdujące się na wybrzeżach (a zatem w pobliżu piasku) są nagie (rysunek 6.15). Innym przykładem jest woda, która robi się coraz jaśniejsza, w miarę zbliżania się do lądu, czyli kiedy zmniejsza się jej głębokość (rysunek 6.16). Przy samej granicy z lądem można nawet dostrzec coś w rodzaju piany morskiej.



rys. 6.16. Woda robi się wyraźnie płytsza w okolicy lądu. Tam, gdzie lądu nie ma, woda jest ciemniejsza.



rys. 6.17. Przykładowy obraz, na którym można zaobserwować zjawiska przenikania się terenów i wygładzania ich granic. Cienkie odnogi wychodzące z rzeki ławy są niemal niewidoczne, zaś te grubsze są wyraźnie ciemniejsze niż główna część terenu. Po lewej, tam, gdzie skała nachodzi na rzekę ławy, widać, jak oba typy terenów się mieszają

6.4.4. Odwzorowanie narysowanych kształtów

Skrawki terenu na obrazie wyjściowym mają kształty bardzo zbliżone do tych narysowanych przez użytkownika na obrazie wejściowym. Można zaobserwować subtelne różnice, polegające na wygładzeniu ostrych krawędzi, a w niektórych przypadkach usunięciu nagłych nieregularności. Małe fragmenty jednego terenu zachodzące na drugi mają tendencję do mieszania się z nim, tracąc swoje kształty.

Efekt wygładzenia jest spodziewany - na obrazach stanowiących dane treningowe poszczególne tereny nie urywają się gwałtownie (z piksela na piksel). Zamiast tego, tereny na swoich granicach stopniowo zanikają albo przechodzą w inne, w sposób organiczny (rysunek 6.17). Ta właściwość została zaobserwowana w przypadku wszystkich wytrenowanych modeli.

6.4.5. Spójność

W przypadku modeli wytrenowanych na domyślnych parametrach, tereny tworzone przy każdym uruchomieniu generatora zawsze wyglądają całkowicie spójnie. Spójność jest tak duża, że jeśli dwie wygenerowane mapy ustawimy obok siebie w taki sposób, by stykały się tym samym typem powierzchni, to trudno jest wskazać granicę między dwoma obrazami. Wskazówkę mogą stanowić jedynie jaśniejsze albo ciemniejsze (w zależności od modelu) linie, które często zostają wytworzone przy granicach obrazów.



rys. 6.18. Przykładowe duże obrazy o wymiarach 512x512. Każdy z nich powstał przez połączenie czterech mniejszych obrazów o wymiarach 512x512. Granice między mniejszymi obrazami są dostrzegalne, jednak niezbyt wyraźne

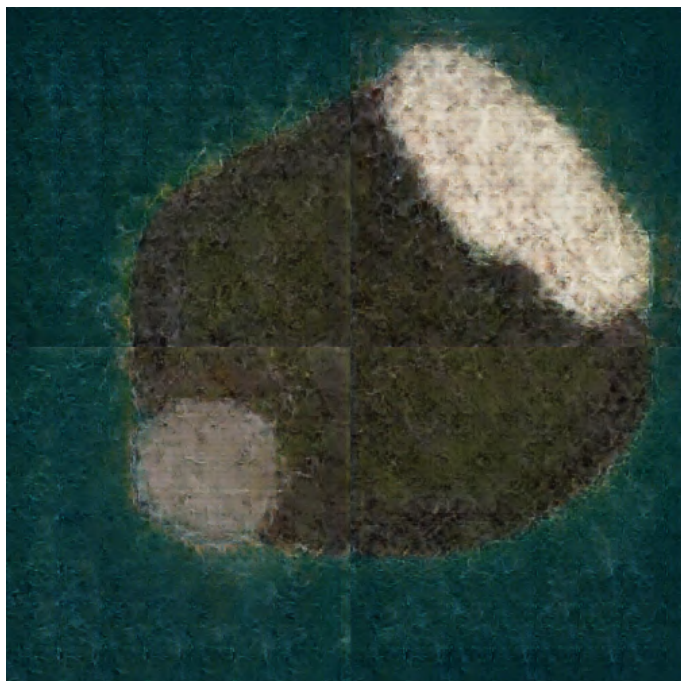
6.4.6. Duże mapy

Korzystając z powyższej właściwości, jesteśmy w stanie tworzyć mapy dużo większe niż 256x256 pikseli. W tym celu tworzymy duże obrazy wejściowe (o bokach których długości to wielokrotności 256), następnie tnijemy je na mniejsze fragmenty (256x256), które po kolei umieszczamy na wejściu generatora. Obrazy wyjściowe, po złożeniu, skutkują powstaniem spójnej mapy (rysunek 6.18). Sytuacja robi się bardziej skomplikowana w przypadku bardziej skomplikowanych wersji modeli, w których kontekst pojawienia się terenu wpływa na jego wygląd. Kontekst może różnić się na każdym z obrazów składowych (które przez generator tworzone są zupełnie osobno, bez wiedzy o pozostałych). W związku z tym, bezpośrednio graniczące ze sobą fragmenty terenu mogą znacznie różnić się wyglądem (rysunek 6.19). Ta właściwość utrudnia użytkownikowi tworzenie większych map. Jest to zadanie, w którym prostsze wersje modeli radzą sobie lepiej.

6.4.7. Uniwersalność

Docelową metodą wprowadzania danych dla generatora jest odręczne rysowanie ich przez użytkownika. Wytrenowanie modelu przy użyciu dokładnie takich danych wejściowych nie było jednak możliwe - nie istnieją bowiem mapy z Divinity, którym odpowiadałby jakikolwiek obraz wejściowy, który może zostać w rozsądnym czasie odtworzony przez człowieka przy użyciu myszy. Mapy z Divinity są zdecydowanie bardziej skomplikowane.

Z tego powodu dane treningowe w znaczący sposób różnią się od danych, na których generator pracuje w praktyce, jednak przy zachowaniu identycznego formatu (rysunek 6.20).



rys. 6.19. Przykład na to, w jaki sposób połączenie pewnych typów terenu może sprawić, że komponenty dużej mapy przestaną być spójne: w prawym górnym rogu brakuje mchu, natomiast w lewym dolnym rogu skała jest jaśniejsza.



rys. 6.20. Porównanie obrazu wejściowego wytworzonego przez użytkownika aplikacji (po lewej) oraz utworzonego na bazie mapy zaczerpniętej z Divinity: Original Sin II.

Pomimo istotnych różnic między obrazami wykorzystywanymi w trakcie szkolenia, wytrenowane modele są w stanie generować mapy dobrej jakości na podstawie obrazów wejściowych tworzonych przez użytkowników. Świadczy to o ich uniwersalności.