Tato práce zkoumá nepřátelské vzory ve strojovém učení, konkrétně v oboru klasifikace obrazu. Nejmodernější modely hlubokého učení dokáží rozpoznat vzory lépe než člověk. Nicméně pokud k obrazům přidáme vhodně zvolený šum, můžeme výrazně snížit přesnost daných modelů. Tato práce zkoumá různé metody útoků i obran proti nepřátelským vzorům. Jednu ze současným obran jsme si vybrali, abychom navrhli útok, který ji porazí. Náš útok využívá evoluční algoritmy. Výsledky našich experimentů ukazují, že nepřátelské vzory vytvořené evolucí a vzory vytvořené se znalostí gradientu se výrazně liší. V neposlední řadě zkoumáme, jak jsou naše nepřátelské vstupy přenositelné mezi různými neuronovými sítěmi.