

Vyhledávání multimédií je stále důležitější vzhledem k prudce rostoucímu objemu multimediálního obsahu. Proto je vyvíjeno mnoho nástrojů pro vyhledávání obrázků a videí, které využívají algoritmy modelování vizuální podobnosti pro vyhledávání podobných obrázků nebo tvorbu různých vizualizací. Tím pádem kvalita modelování podobnosti je pro tyto nástroje klíčová. Tato práce zkoumá různé modely podobnosti, jejich shodu s lidskými anotacemi a potenciální zlepšení. Do studie byly zahrnuty 3 třídy modelů podobnosti: modely založených na barvách, SIFTu nebo hlubokých neuronových sítí. Za tímto účelem byla provedena uživatelská studie s cílem vytvořit dataset relativních podobností obrázků s obecnými i specifickými obrázky. V této studii byli účastníci požádáni, aby vždy vybrali mezi dvěma možnostmi tu, která byla podobnější hlavnímu obrázku. Shromážděná data ukázala lepší výsledky modelů založených na hlubokých neuronových sítích ve srovnání s ostatními hodnocenými variantami. Nicméně všechny modely podobnosti si vedly výrazně lépe než náhodný odhad. Abychom dále zvýšili přesnost modelů podobnosti, vyladili jsme model W2VV++ pomocí získaného datasetu. Díky tomu jsme v některých doménách dosáhli výrazného zlepšení.