

ImageCritique – Utvikling av og erfaringer med et KI-basert hjelpemiddel for bildekritikk i undervisning

H. L. Sundbø, K. Mosgren, T. Bergholtz og P. A. Kvamsø
*Institutt for informasjons- og kommunikasjonsteknologi, Fakultet for teknologi og realfag,
Universitetet i Agder*

ABSTRACT: En multimodal KI-chatbot med navnet ImageCritique, ble utviklet for å gi studenter innen multimediefag saklige tilbakemeldinger på bildebaserte arbeider. Chatboten er bygget på Chat GPT og benytter oppgavespesifikke kriterier i systemprompten for å gi målrettede tilbakemeldinger. Verktøyet ble testet i emnene DAT104 Multimediedesign og DAT227 Foto ved Universitetet i Agder. Effekten ble undersøkt gjennom en kontrollert forhåndstest i UiAs Usability Lab (n = 5) og en nettbasert spørreundersøkelse etter ordinær bruk i undervisning (n = 20). Resultatene viser at ImageCritique gjennomgående oppleves som nyttig og læringsstøttende. En sekundær effekt var at lærere sparte tid ved å slippe å skrive skriftlige tilbakemeldinger. Funnene peker mot at presist formulerte instruksjoner er en sentral forutsetning for høy tilbakemeldingskvalitet, og at KI-baserte chatboter har reelt pedagogisk potensiale i teknisk-kreative fag.

1. INNLEDNING

Studenter i teknisk-kreative fag som foto og bildedesign trenger jevnlig og konkret tilbakemelding for å utvikle faglig kompetanse. I praksis er læreres kapasitet til å gi individuelle tilbakemeldinger begrenset, særlig i emner med mange studenter og hyppige innleveringer. Dette skaper et gap mellom studentenes behov for rask respons og hva lærerne faktisk kan levere innenfor normale ressursrammer. Formativ vurdering, tilbakemelding gitt underveis i arbeidsprosessen, er vel dokumentert som en av de sterkeste påvirkningsfaktorene for læring [1], og det er nettopp denne typen tilbakemelding som lettest tapes når kapasiteten er stram.

ImageCritique er utviklet for å fylle dette gapet. Verktøyet er en multimodal KI-chatbot som gir studenter konstruktiv kritikk på bildebaserte arbeider, umiddelbart og tilgjengelig når som helst. I fotofag retter tilbakemeldingen seg mot komposisjon og kameratekniske innstillinger, mens chatboten i bildedesign i større grad fokuserer på om formelle oppgavekrav er oppfylt.

Denne artikkelen beskriver utviklingen av ImageCritique, metodene som ble brukt for å evaluere det, og erfaringene fra bruk i undervisning ved Universitetet i Agder (UiA). Arbeidet er gjennomført som en del av Program for utdanningskvalitet (PRUK) ved UiA i perioden 2024/2025.

2. UTVIKLING AV IMAGECRITIQUE

ImageCritique ble utformet som en egen chatbot utviklet med ChatGPT. Grunnideen er enkel: studenten laster opp et bilde, stiller et spørsmål eller ber om en vurdering, og chatboten svarer med faglig begrunnede tilbakemeldinger basert på oppgavens spesifikke kriterier.

Kriteriene er bygget inn i systemprompten: en forhåndsdefinert instruksjonstekst som styrer chatbotens svar. Kvaliteten på denne instruksjonen er avgjørende, språkmodeller er høyst sensitive for prompt-utformingen, og presise instruksjoner gir vesentlig bedre og mer målrettede svar enn vage eller generelle formuleringer [2]. Under utviklingen ble ulike formuleringer testet og justert frem til tilbakemeldingene konsekvent traff oppgavens faglige krav.

Flere KI-plattformer ble vurdert og testet i utviklingsprosessen, blant annet ChatGPT, Microsoft Copilot Studio og Dify. Nyere resoneringsmodeller ga mer presise og nyanserte tilbakemeldinger enn eldre modeller, og valget falt til slutt på ChatGPTs API som teknisk fundament.

3. METODER

For å evaluere ImageCritique ble to metoder benyttet i kombinasjon: en kontrollert forhåndstest og en spørreundersøkelse etter ordinær bruk.

Før verktøyet ble tatt i bruk i ordinær undervisning, ble det gjennomført en kontrollert test i UiAs Usability Lab med fem frivillige studenter (n = 5). Formålet var å avdekke styrker og svakheter i både grensesnittet og tilbakemeldingskvaliteten, mens det fortsatt var mulig å gjøre justeringer.

Etter normal bruk i emnene DAT104 Multimediedesign og DAT227 Foto ble det gjennomført en nettbasert spørreundersøkelse med 20 respondenter. Spørreundersøkelsen bestod av lukkede spørsmål med forhåndsbestemte svaralternativer, samt åpne kommentarfelt. Spørsmålene dekket fire områder: (1) responstid og tilgjengelighet, (2) tilbakemeldingenes nytte og presisjon, (3) læringseffekt sammenlignet med tradisjonell veiledning, og (4) ønske om fremtidig bruk.

4. RESULTATER

Tilbakemeldingene fra brukerne var overveiende positive, og verktøyet ble raskt populært. 85 prosent vurderte responstiden som akseptabel eller bedre. Nytteverdien var mer variert: 28 prosent beskrev tilbakemeldingene som «generelt relevante og nyttige», mens 44 prosent svarte «varierende». Presisjon var en klar svakhet – 53 prosent opplevde tilbakemeldingene som «litt uklare», og åpne svar pekte på inkonsistens: «Ulike svar på samme bilde hvis det ble levert flere ganger. Veldig varierende score.»

På spørsmålet om ImageCritique bidro til mer effektiv læring enn en lærerkommentar, svarte ingen «i høy grad». 35 prosent svarte «i noen grad» og 30 prosent «hverken/eller». ImageCritique oppleves altså ikke som et likeverdig alternativ til lærertilbakemelding, men snarere som et nyttig supplement.

Resultatene skilte seg tydelig mellom de to emnene. I DAT104 Multimediedesign, med klart definerte oppgavekrav, opplevde studentene verktøyet som en konkret hjelp til å sjekke hva som ikke var oppfylt: «På kreative oppgaver som *Space Hunters*¹ er den bra. Men på oppgaver som *Kinalykter*² kunne den ikke se nøye nok om jeg hadde gjort feil.» 47 prosent oppga at de ønsker KI-tilbakemelding enten alene (6 %) eller kombinert med veileder (41 %). I DAT227 Foto var bildet mer blandet: bare 15 prosent ønsket mer KI-basert tilbakemelding, og studentene etterlyste mer konstruktiv kritikk: «Den ga litt tips men veldig positivt og ønsker mer konstruktiv kritikk.»

¹ *Space Hunters* er en oppgave i Adobe Photoshop hvor 5 bilder skal settes sammen til en utenomjordisk scene med gitte forutsetninger.

² *Kinalykter* er en oppgave i Adobe Photoshop hvor elementer i et fotografi skal selekteres og manipuleres med hensyn til farger.

5. DISKUSJON

Funnene peker på at ImageCritique fungerer best der oppgavekravene er klart definerte og målbare. Når kriteriene er eksplisitte, kan en presis systemprompt gi direkte faglig gevinst. I emner med større faglig og kunstnerisk skjønn er forholdet mellom KI-tilbakemelding og studentens autonomi mer komplisert. Hattie og Timperley [1] påpeker at effektiv tilbakemelding må svare på tre spørsmål: hvor er jeg nå, hvor skal jeg, og hvordan kommer jeg dit. ImageCritique klarer dette rimelig godt for oppgavetyper med klare målkriterier, men sliter med å gi meningsfull «vei videre»-tilbakemelding i mer åpne, kunstneriske kontekster.

Tendensen til *for* positive tilbakemeldinger er en kjent svakhet ved store språkmodeller [2], [3], og kan motvirkes gjennom eksplisitte instruksjoner i systemprompten om å gi balanserte og kritiske vurderinger. Det åpne svaret «Gjør den mer konsis» peker i samme retning: tydelighet og korthet bør vektlegges i videre promptutvikling.

Den strukturelle utfordringen er verdt å nevne eksplisitt: ImageCritique er per nå basert på ChatGPT, som ikke er en godkjent KI-plattform ved UiA. Det er derfor ikke mulig å innføre verktøyet som et obligatorisk element i undervisningen, og det vil ikke være aktuelt å benytte det til vurdering av innleverte studentarbeider. En overgang til en institusjonelt godkjent plattform er en forutsetning for bredere bruk.

6. OPPSUMMERING

ImageCritique demonstrerer at KI-baserte multimodale chatboter kan ha reelt pedagogisk potensiale som læringsstøttende verktøy i teknisk-kreative fag. Verktøyet gir studentene umiddelbare tilbakemeldinger og frigjør lærernes tid. Effekten er størst i emner med klart definerte oppgavekrav, der presise instruksjoner i systemprompten er avgørende for kvaliteten.

For å realisere det fulle potensialet bør fremtidig utvikling rette seg mot to områder: bedre balanse i tilbakemeldingenes tone, mindre positiv, mer konkret og konsis, og en overgang til en institusjonelt godkjent KI-plattform. Erfaringene fra dette prosjektet gir et godt grunnlag for begge deler.

REFERANSER

- [1] J. Hattie and H. Timperley, "The Power of Feedback," *Review of Educational Research*, vol. 77, no. 1, pp. 81–112, 2007, doi: 10.3102/003465430298487.
- [2] E. Kasneci et al., "ChatGPT for good? On opportunities and challenges of large language models for education," *Learning and Individual Differences*, vol. 103, p. 102274, 2023, doi: 10.1016/j.lindif.2023.102274.
- [3] J. Rudolph, S. Tan, and S. Tan, "ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?" *Journal of Applied Learning and Teaching*, vol. 6, no. 1, pp. 342–363, 2023, doi: 10.37074/jalt.2023.6.1.9.