Emvi, een herkenningsapp met AI

Louis Pool, Koninklijke Visio



Emvi.ai of kortweg EMVI is een app waarmee je een foto maakt, waarna de app beschrijft wat er op de foto staat. Vervolgens kun je de app vragen stellen over de inhoud van de foto. Zo kun je de app gebruiken om teksten te laten voorlezen, de ruimte waarin je je bevindt te beschrijven, een product te omschrijven of je te helpen zoeken als je iets verloren hebt.

EMVI is toegankelijk voor slechtziende mensen die nood hebben aan goede contrasten en duidelijke knoppen en voor blinde mensen die de VoiceOver schermlezer gebruiken op hun iPhone of TalkBack op hun Androidsmartphone. De app kost in Nederland 4,99 euro per maand en in België 49,99 euro per maand. Voor levenslang gebruik kun je ook eenmalig €200 betalen. Het gebruik en de functionaliteit van de app is identiek voor de Apple- en de Android-versies.

# Waar kun je de EMVI app voor gebruiken?

De EMVI-app is veelzijdig en biedt een ruim aantal mogelijkheden. We doen een greep uit het ruime aanbod van functies:

* Een omschrijving krijgen van een foto.
* Een beschrijving krijgen van een product op basis van de streepjescode.
* Iets vinden dat je verloren hebt.
* Ontdekken wat er op een rommelige tafel ligt.
* Informatie in een agenda of weekplanner bekijken.
* Tekst van papier laten voorlezen, zoals een brief, een folder of een informatiebordje.
* Informatie krijgen over kunstwerken of bekende gebouwen.
* Kleuren van een voorwerp aangeven.
* Hulp bieden bij het lopen van een traject.
* Een QR-code scannen en laten vertellen wat er op de site staat waar de QR-code naar verwijst.

Dankzij de EMVI-app heb je snel en eenvoudig toegang tot de informatie die je wenst.

# Wat is het verschil tussen EMVI en andere herkenningsapps?

Er zijn zowel zaken die EMVI beter kan, als dingen waar andere apps beter in zijn:

* EMVI is uitgerust met een grote microfoonknop. Dit maakt dat je snel en gemakkelijk vragen stellen over de foto die je hebt gemaakt. Vergeleken met bijvoorbeeld Be My Eyes kom je zo sneller tot de informatie die je zoekt.
* EMVI heeft ook nadelen ten opzichte van andere apps. Waar apps als Seeing AI en Envision je helpen om een foto te maken van een tekst moet je bij EMVI zonder begeleiding zelf de foto maken. Dat kan lastig zijn, zeker wanneer je niet over restvisus beschikt. In dat geval moet je waarschijnlijk oefenen hoe ver je de telefoon van een tekst moet houden om een foto te maken. Gelukkig zijn hier wel trucs voor. Je vind ze in het artikel: [Lees meer over blind een document fotograferen](https://kennisportaal.visio.org/documenten/blind-een-document-fotograferen/).
* Wat EMVI, in tegenstelling tot andere herkenningsapps, niet doet is de tekst tonen die de spraakassistent voorleest. Je kunt de tekst die je hebt gefotografeerd dus ook niet doorsturen of nalezen.
* In de beschrijving van een foto, schilderij of de omgeving richt EMVI zich meer dan andere apps op de sfeer en het gevoel dat iets uitstraalt. Zo zal EMVI bijvoorbeeld zeggen: De vrouw op het schilderij staart naar de horizon. De bergpieken op de achtergrond, samen met de donkere aardse kleuren zorgen voor een nostalgisch gevoel.

# Aan de slag met EMVI

Wanneer je de app opent en je bent ingelogd, word je eerst welkom geheten door de spraakassistent van de app. Centraal in beeld verschijnt er een grote, ronde, blauwe knop met daarop de afbeelding van een camera. Je bent dan in de fotofunctie terechtgekomen. Onderaan het scherm staat een knoppenbalk met de vier hoofdfuncties. Je selecteert een van die functies niet door op een van de vier knopjes te tikken, maar met korte horizontale veegjes naar links of rechts (om naar de volgende of vorige functie te gaan). Met de VoiceOver schermlezer op een iPhone gebruik je drie vingers voor die horizontale veegjes; met de TalkBack schermlezer op een Android-telefoon gebruik je twee vingers. Op die manier maak je een keuze uit de vier functies van de app:

* De Fotofunctie, een grote blauwe knop met een camera-pictogram (de telefoon geeft een voelbaar tikje wanneer je op die functie komt) om een foto te nemen.
* De Leesfunctie, een grote blauwe knop met pictogram van een hand die naar een opengeslagen boek wijst (de telefoon geeft twee voelbare tikjes wanneer je op die functie komt), waarmee je tekst kunt lezen.
* De Scanfunctie, een grote blauwe knop met een QR-scan-pictogram (de telefoon geeft drie voelbare tikjes wanneer je op die functie komt) waarmee je QR- of streepjescodes kunt scannen.
* De Spraakingavepagina, een grote blauwe knop met een microfoon-pictogram, waarmee je allerlei vragen en vervolgvragen kunt stellen.

We overlopen die vier functieknoppen en wat je ermee kunt doen.

## 1. De fotofunctie

Wanneer je de fotofunctie activeert, verschijnt het camerabeeld op je scherm en hoor je de spraakassistent aftellen van drie naar een. Het is belangrijk dat je de camera gericht hebt én stil houdt tegen het eind van de aftelling, want meteen daarna wordt er een foto gemaakt. Vervolgens verschijnt de knop van de spraakingavepagina. Er wordt nog niets verteld over de foto, want je moet eerst met een vraag aangeven welke informatie je uit de foto wilt halen. De gesproken feedback zal afhankelijk zijn van hoe jij je vraag formuleert.

## 2. De spraakingavepagina

Op de spraakingavepagina zie je de ronde knop met daarop een microfoon. Wanneer je die activeert kun je aan de spraakassistent een vraag stellen over de foto die je gemaakt hebt. Ben je klaar met je vraag, dan moet je nogmaals op die knop drukken. Zo weet de app dat je klaar bent met je vraag te stellen. Het zou handig zijn moest de app zelf detecteren wanneer je klaar bent met je vraag, zoals dat bij Siri of Google Assistent werkt. De spraakassistent spreekt vervolgens vrijwel direct een antwoord uit (het antwoord verschijnt niet op het scherm). De feedback is doorgaans zeer gedetailleerd. Kenmerkend is dat er veel aandacht wordt besteed aan de sfeer die een foto uitstraalt en aan de emoties die te zien zijn bij personen op de foto. De knop met de microfoon blijft op het scherm staan. Zo kun je een vervolgvraag stellen.

Een voorbeeld: je fotografeert een keukenaanrecht en stelt daarna de vrij algemene vraag: “Kun je de foto beschrijven?”. EMVI beschrijft wat er te zien is en vermeldt onder andere dat er drie paprika’s te zien zijn. Daarna stel je de vervolgvraag: “Welke kleur hebben de paprika’s?”. EMVI geeft aan dat er twee rode en een groene paprika op het aanrecht liggen. Vervolgens kun je nog meerdere extra vervolgvragen stellen.

Doordat je zelf een vraag kunt formuleren, kun je de antwoorden van EMVI zelf in de gewenste richting sturen. Een paar voorbeelden van vragen:

* “Hoeveel personen zie je in de kamer?”
* “Zien de personen er gelukkig uit?”
* “Kun je iets meer zeggen over het gebouw dat in beeld is?”
* “Van welke schilder is het schilderij?”
* “Zijn er obstakels op de stoep voor mij?”
* “Is er een zebrapad in beeld en kun je aangeven hoe ik daar naartoe stap?”
* Als je iets verloren hebt gelegd: “Is mijn sleutelbos te zien in het beeld?”
* “Brandt de verlichting in deze ruimte?”
* “Hoe laat is het op de klok die je in deze ruimte ziet?”
* “Beschrijf het voorwerp op de tafel en geef daarbij vooral de kleuren van het voorwerp aan.”
* …

We hebben bijvoorbeeld met de laatste vraag ondervonden dat deze app veel betere opties biedt om kleuren aan te geven dan de specifieke kleurdetectie-functies van andere AI-apps of kleurdetectie-apps.

Je kunt overigens ook vragen stellen zonder een foto te nemen. Bijvoorbeeld:

* “Gooi twee dobbelstenen en spreek het totaal aantal ogen uit.”
* “Welke veegbeweging moet ik maken om met VoiceOver de rotor te bedienen?”
* “Bij welke organisatie kan ik in Vlaanderen of Nederland terecht voor hulpverlening aan blinde mensen?”
* …

Het interessante aan het conversationele model van de app is dat je naar hartenlust kunt experimenteren met om het even welke vraag die je kunt bedenken. En zo ga je ongetwijfeld nieuwe en mogelijks ook erg nuttige functies ontdekken.

Als je een nieuwe foto wilt maken, dan ga je met horizontale veegjes opzij (met een, twee of drie vingers) tot je bij de fotofunctie bent.

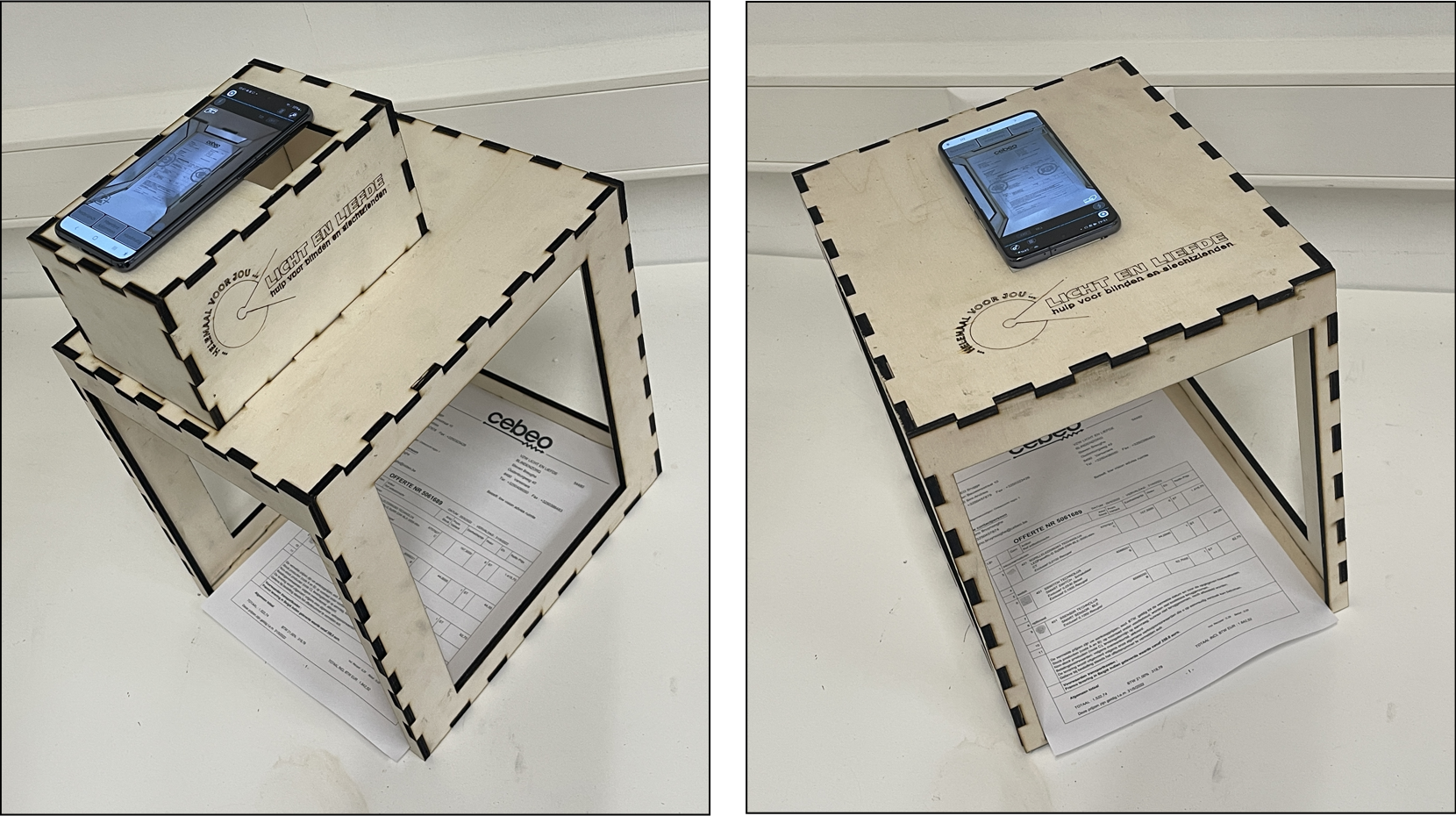
## 3. De leesfunctie

Met de leesfunctie kun je tekst laten voorlezen: tekst van een gedrukt document, tekst van een verpakking, tekst op een lcd-scherm van een toestel … Richt je telefoon op een stuk tekst en druk op de knop van de leesfunctie. Net zoals bij de fotofunctie zal de telefoon aftellen van drie naar een en daarna de foto maken. De tekst op de foto wordt vervolgens direct voorgelezen.

Tijdens het voorlezen verschijnt de spraakingavepagina met de microfoonknop. Die biedt je de mogelijkheid om specifieke vragen te stellen over de voorgelezen tekst, zoals:

* “Kun je een beknopte samenvatting maken van het document?”
* “Kun je de tweede alinea nog eens herhalen?”
* “Van wie is deze brief afkomstig?”
* “Kun je het factuurbedrag en de uiterste betaaldatum nog eens voorlezen?”
* Als je een anderstalig document hebt gescand: “Vertaal de tekst naar het Nederlands.”
* …

EMVI biedt geen hulp bij het blind nemen van een foto (zoals bijvoorbeeld Seeing AI of OneStep Reader, die aangeven wanneer het volledige document in beeld is en dan automatisch de foto nemen). Dat is jammer. Als je vaak documenten leest die je blind fotografeert met EMVI kunnen we daar twee tips voor geven. Je kunt het artikel “[Blind een document fotograferen](https://kennisportaal.visio.org/documenten/blind-een-document-fotograferen/)” op het Kennisportaal eens doornemen. Daarnaast kun je overwegen om een scanstandaard te gebruiken. Die zorgt er altijd voor dat het hele document in beeld is en dat je de camera niet beweegt tijdens het nemen van de foto.



## 4. De scanfunctie

Met de scanfunctie scan je eenvoudig een QR- of streepjescode die naar een website leidt. Nadat je de scanfunctie hebt geactiveerd richt je de camera op de code. Die wordt direct gescand, waarna EMVI de achterliggende website bekijkt en nuttige informatie uit die website op een begrijpelijke manier verwoordt.

Er zijn twee mogelijkheden:

* Je hebt een QR-code gescand: EMVI vat de informatie op de gekoppelde website voor jou samen.
* Je hebt de streepjescode op de verpakking van een product (voeding, drank, hygiënisch product, medicatie …) gescand: EMVI zoekt de productinformatie en leest die in mensentaal voor. De beschrijving die EMVI geeft van het product is veel informatiever dan bij de meeste andere streepjescode-scanners. Zoals andere streepjescode-scanners kent ook EMVI niet alle producten.

Na het scannen van de QR- of streepjescode verschijnt de spraakingavepagina, waarmee gesuggereerd wordt dat je nog bijkomende vragen kunt stellen bij de voorgelezen informatie. Maar dat lukte ons niet omdat de app telkens de melding gaf dat er geen foto genomen werd.

## 5. Opgeslagen foto’s laten beschrijven

Je kunt EMVI ook vragen om een foto uit je fotobibliotheek of uit een app zoals bijvoorbeeld WhatsApp te beschrijven. Ga daarvoor naar de app met de foto die je wilt laten beschrijven. Je selecteert de foto en activeert de knop ‘Delen’ en je kiest vervolgens voor ‘EMVI’. Daarna ga je naar de spraakingavepagina van de EMVI-app, activeer je de microfoon en stel je een vraag over de foto.

# PhoneCam

De ontwikkelaars van EMVI werken erg hard aan de app. We zullen geregeld nieuwe versies zien verschijnen en er zitten heel wat mooie nieuwe functies in de pipeline. Van een van die toekomstige nieuwigheden mogen we alvast een tipje van de sluier lichten: de PhoneCam.



PhoneCam is een klein (1,5 bij 5 bij 2 cm) cameraatje met een geïntegreerde microfoon dat je aan je kledij bevestigt en dat de beelden doorstuurt naar de gekoppelde smartphone. De makers van EMVI zijn in hun app ondersteuning aan het ontwikkelen voor dat cameraatje.

We hebben al even de kans gehad om een dag rond te lopen met de PhoneCam en mochten ervaren dat die camera de EMVI-app naar een hoger niveau tilt. Je kunt EMVI namelijk gebruiken terwijl je telefoon in je jaszak blijft zitten. Het cameraatje heeft een drukknopje dat je indrukt om een foto te nemen, vervolgens nogmaals indrukt om een vraag te stellen en een derde maal indrukt om de vraag te beëindigen. Vervolgens geeft EMVI de feedback zoals hierboven beschreven. Je kunt vervolgvragen stellen door het knopje nogmaals in te drukken. Als je een nieuwe foto wilt nemen, druk je tweemaal na elkaar op het knopje en kun je heel het bedieningsproces opnieuw doorlopen. We gaan verder niet te veel in op details want deze nieuwigheid is in volle ontwikkeling.

Het was een veelbelovende ervaring en we kijken er naar uit om dat op de markt te zien komen. Tot slot het goede nieuws: in tegenstelling tot de dure camerabrillen, kost dat toestelletje nog geen 200 euro.

# Heb je nog vragen?

Mail naar [kennisportaal@visio.org](mailto:kennisportaal@visio.org), of bel [088 585 56 66](tel:0885855666)

Meer artikelen, video’s en podcasts vind je op [kennisportaal.visio.org](https://kennisportaal.visio.org/)

Koninklijke Visio

expertisecentrum voor slechtziende en blinde mensen

[www.visio.org](http://www.visio.org)